Lean Six Sigma Green Belt

by
Rajiv Purkayastha
Six sigma MBB

Analyze Phase

Analyze Phase Overview

Why is the Analyze phase important?

This phase is important because it clearly defines how well the process is currently performing and identifies how much the process will be improved.

Cause & Effect Matrix

A Tool to Identify and Quantify Sources of Variation

- Relates the key inputs to the key outputs (customer requirements) using the process map as the primary information source
- Key outputs are scored as to importance to the customer
- Key inputs are scored as to relationship to key outputs
- Pareto of key inputs to evaluate in the **FMEA** and control plans
- Input into the initial evaluation of the Process Control Plan

Cause & Effect Matrix Steps

- Identify key customer requirements (outputs) from process map or other sources
- Rank order and assign priority factor to each Output (usually on a 1 to 10 scale)
- Identify all process steps and materials (inputs) from the Process Map
- Evaluate correlation of each input to each output
  - *low score*: changes in the input variable (amount, quality, etc.) have small effect on output variable
  - *high score*: changes in the input variable can greatly affect the output variable
- Cross multiply correlation values with priority factors and sum for each input

Examples

Bottle Production Big Block Diagram

Bottle Production

Upstream	**Inputs**	**Outputs**	**Downstream**
Resin IV	Regrind %	Break % at Molder	Break % at Filler
Other resin Properties	Barrell Temperature	Bottle Output	Etc…
Bottle Design	Screw Speed	Melt Temperature	
Etc…	Screw Design	Clarity	
	Etc…	Weight	
		Wall Thickness Variation	
		Etc…	

Note: Only a partial list

Bottle Production Example

Bottle Production Block Steps Diagram

Inputs	Process step	Outputs
Resin	Vacuum out of box	Pellet change
	Blending with Regrind	
	Dryer Hopper (Regrind + virgin)	
	Vacuum into machine hopper	
Resin Regrind%-C Barrel temperature-C Screw speed-C Meter cooling temp-C Screw design-C Barrel/screw condition-U	Melt Resin	Melt Temperature Melt strength MW/IV Throughput
Parison programming-C Head adapter design-C Head design (PVC, HDPE)-C Tooling Zone temps Die tip temp Support air (straight and controlled) Parison cut	Extrude Parison	Parison length Parison diameter Parison temp Thickness Profile Clarity Melt stability

Note: Only a partial Map

Cause & Effect Matrix Form

1. List Key Outputs

		ting of rtance to stomer																
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
			Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Total
	Process Step	**Process Input**																
1																		0
2																		0
3																		0
4																		0
5																		0
6																		0
7																		0
8																		0

Bottle Production Example

		Rating of Importance to Customer							
			1	2	3	4	5	6	
			Break %	Melt Temp	Bottle Output	Clarity	Weight	Wall Variation	Total
	Process Step	Process Input							
									0
									0
									0
									0
									0
									0
									0
									0
									0
									0
									0
									0
									0

1. List Key Outputs

The Outputs are defined in Step 1 of Process Mapping

Cause & Effect Matrix Form

		Rating of Importance to Customer																
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
			Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement			
	Process Step	Process Input																
1																		0
2																		0
3																		0
4																		0
5																		0
6																		0
7																		0
8																		0

2. Rank Outputs as to Customer importance

Bottle Production Example

2. Rank Outputs as to Customer importance

	Rating of Importance to Customer	10	1	9	9	5	8	
		1	2	3	4	5	6	
		Break %	Melt Temp	Bottle Output	Clarity	Weight	Wall Variation	Total
Process Step	**Process Input**							

Cause & Effect Matrix Form

3. List Key Inputs by Process Step

		Rating of Importance to Customer																
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
			Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Total
	Process Step	**Process Input**																
1																		0
2																		0
3																		0
4																		0
5																		0
6																		0
7																		0
8																		0

Note: Information obtained from process map

Bottle Production Example

Note: Only a partial list of inputs

	In						
				9	9	5	8
				3	4	5	6
		Break	Melt Te	Bottle Output	Clarity	Weight	Wall Variation
Process Step	Process Input						
Melt Resin	Resin						
Melt Resin	Barrell Temp						
Melt Resin	Screw Speed						
Melt Resin	Screw Design						
Melt Resin	Regrind%						
Melt Resin	Barell/Screw Condition						
Melt Resin	Screw Tip Cooling						
Extrude Parison	Programing						
Extrude Parison	Die Tip Temp						
Extrude Parison	Head Design						
Extrude Parison	Tooling						

3. List Key Inputs by Process Step

This step uses the Process Map inputs directly. Notice the Process Inputs follow the Process map step-by-step.

Cause & Effect Matrix Form

4. Relate Inputs to Outputs

		Rating of Importance to Customer																
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
			Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Total
	Process Step	**Process Input**																
1																		0
2																		0
3																		0
4																		0
5																		0
6																		0
7																		0
8																		0

Relating Inputs to Customer Requirements

- You are ready to relate the customer requirements to the process input variables
- ***Correlational scores***: No more than 4 levels
  - 0, 1, 3 and 9
- Assignment of the scoring takes the most time
- To avoid this, spell out the criteria for each score:
  - ***0*** = No correlation
  - ***1*** = The process input only remotely affects the customer requirement
  - ***3*** = The process input has a moderate effect on the customer requirement
  - ***9*** = The process input has a direct and strong effect on the customer requirement

Bottle Production Example

4. Relate Inputs to Outputs

		Rating of Importance to Customer	10	1	9	9	5	8	
			1	2	3	4	5	6	
			Break %	Melt Temp	Bottle Output	Clarity	Weight	Wall Variation	Total
	Process Step	**Process Input**							
1	Melt Resin	Resin	9	9	3	9	9	9	324
1	Melt Resin	Barrell Temp	3	9	9	1	3	3	168
1	Melt Resin	Screw Speed	3	9	9	1	3	3	168
1	Melt Resin	Screw Design	3	9	9	1	1	1	142
1	Melt Resin	Regrind%	3	1	1	3	3	3	106
1	Melt Resin	Barell/Screw Condition	3	3	3	1	1	1	82
1	Melt Resin	Screw Tip Cooling	1	1	3	0	3	3	77
2	Extrude Parison	Programing	3	3	9	0	9	9	231
2	Extrude Parison	Die Tip Temp	3	3	3	9	3	9	228
2	Extrude Parison	Head Design	3	9	3	3	3	9	180
2	Extrude Parison	Tooling	3	3	3	3	3	9	174
2	Extrude Parison	Support Air	1	0	9	0	1	1	104
2	Extrude Parison	Lower Manifold	3	3	3	3	1	1	100

Note: Only a partial list of inputs

This is a subjective estimate of how influential the key inputs are on the key outputs

Cause & Effect Matrix Form

		Rating of Importance to Customer																
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
			Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement	Requirement				
	Process Step	Process Input																
1																		0
2																		0
3																		0
4																		0
5																		0
6																		0
7																		0
8																		0

5. Cross-multiply and prioritize

Sum of (Rating x Correlation Score) values for all Requirements

Bottle Production Example

5. Cross-multiply and prioritize

		Rating of Importance to Customer	10	1	9	9	5	8	
			1	2	3	4	5	6	
			Break %	Melt Temp	Bottle Output	Clarity	Weight	Wall Variation	Total
	Process Step	**Process Input**							
1	Melt Resin	Resin	9	9	3	9	9	9	324
1	Melt Resin	Barrell Temp	3	9	9	1	3	3	168
1	Melt Resin	Screw Speed	3	9	9	1	3	3	168
1	Melt Resin	Screw Design	3	9	9	1	1	1	142
1	Melt Resin	Regrind%	3	1	1	3	3	3	106
1	Melt Resin	Barell/Screw Condition	3	3	3	1	1	1	82
1	Melt Resin	Screw Tip Cooling	1	1	3	0	3	3	77
2	Extrude Parison	Programing	3	3	9	0	9	9	231
2	Extrude Parison	Die Tip Temp	3	3	3	9	3	9	228
2	Extrude Parison	Head Design	3	9	3	3	3	9	180
2	Extrude Parison	Tooling	3	3	3	3	3	9	174
2	Extrude Parison	Support Air	1	0	9	0	1	1	104
2	Extrude Parison	Lower Manifold	3	3	3	3	1	1	100

Note: Only a partial list of inputs

We now start getting a feel for which variables are most important to explaining variation in the outputs

Bottle Production Example

	Rating of Importance to Customer	10	1	9	9	5	8	
		1	2	3	4	5	6	
		Break %	Melt Temp	Bottle Output	Clarity	Weight	Wall Variation	Total
Process Step	Process Input							
Melt Resin	**Resin**	**9**	**9**	**3**	**9**	**9**	**9**	**324**
Blow Bottle	**Mold Design**	**9**	**0**	**9**	**9**	**0**	**9**	**324**
Extrude Parison	Programing	3	3	9	0	9	9	231
Extrude Parison	Die Tip Temp	3	3	3	9	3	9	228
Blow Bottle	**Mold Water Temp**	**9**	**0**	**9**	**3**	**0**	**0**	**198**
Blow Bottle	**Water Volume/Cooling Rate**	**9**	**0**	**9**	**3**	**0**	**0**	**198**
Extrude Parison	Head Design	3	9	3	3	3	9	180
Extrude Parison	Tooling	3	3	3	3	3	9	174
Blow Bottle	**Pinch Design**	**9**	**0**	**9**	**0**	**0**	**0**	**171**
Melt Resin	Barrell Temp	3	9	9	1	3	3	168
Melt Resin	Screw Speed	3	9	9	1	3	3	168
Melt Resin	Screw Design	3	9	9	1	1	1	142
Blow Bottle	**# of Mold Cooling Zones**	**9**	**0**	**3**	**0**	**0**	**0**	**117**
Tail Detab	**Time from extraction to Detab**	**9**	**0**	**3**	**0**	**0**	**0**	**117**

We have sorted on the cross-multiplied numbers and find that the input variables in the box above are the most important

We can now evaluate the control plans for these input variables

Note: Only a partial list of inputs

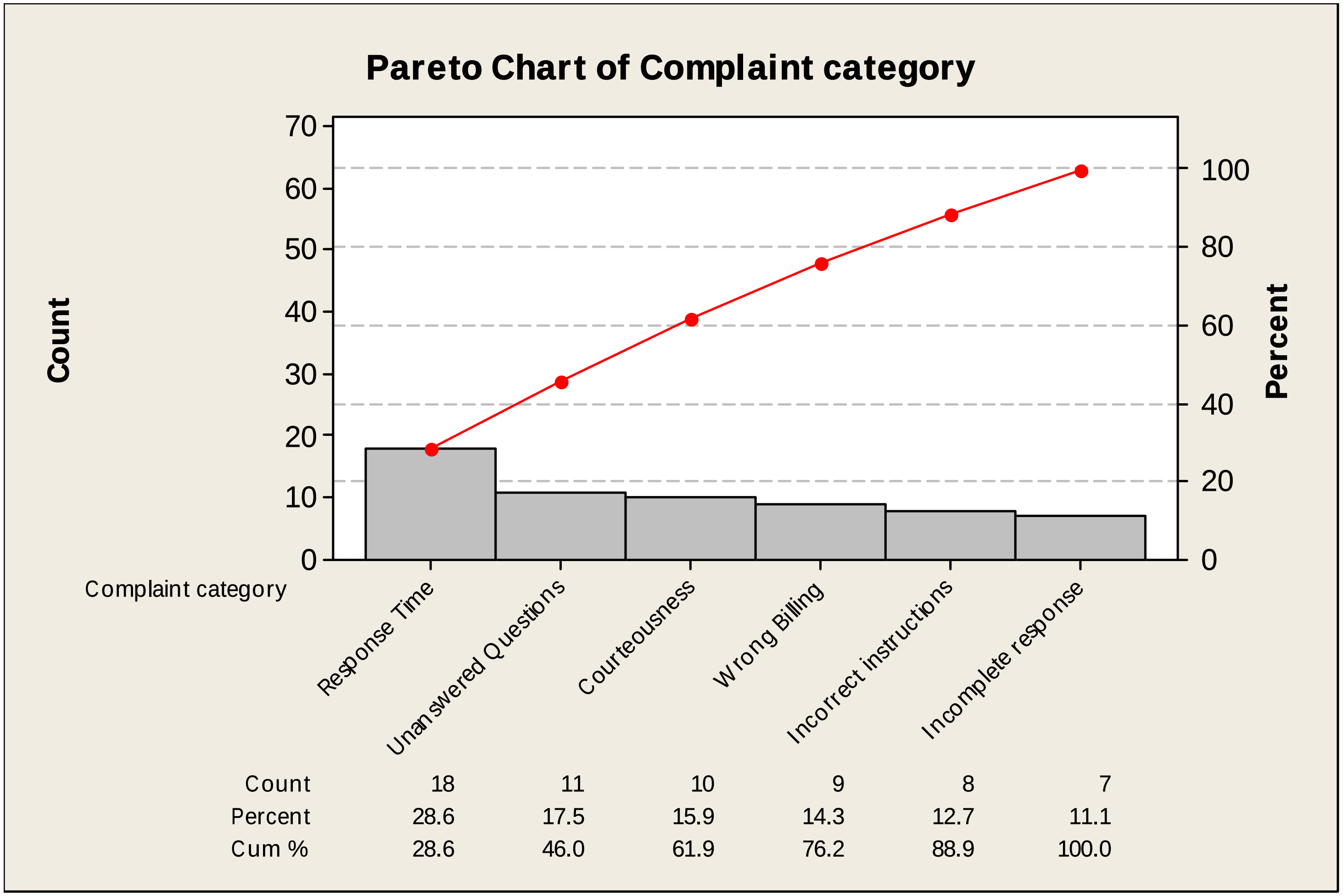
Pareto Analysis

Pareto Chart

- The Pareto chart is named for an Italian economist who found that that the largest part of the Italian wealth was held by a very small percentage of people in the course his analysis he developed a graphic method for displaying the relative importance of causes or factors.

- The Pareto principle or 80/20 rule : 80% of results come from 20% of the causes

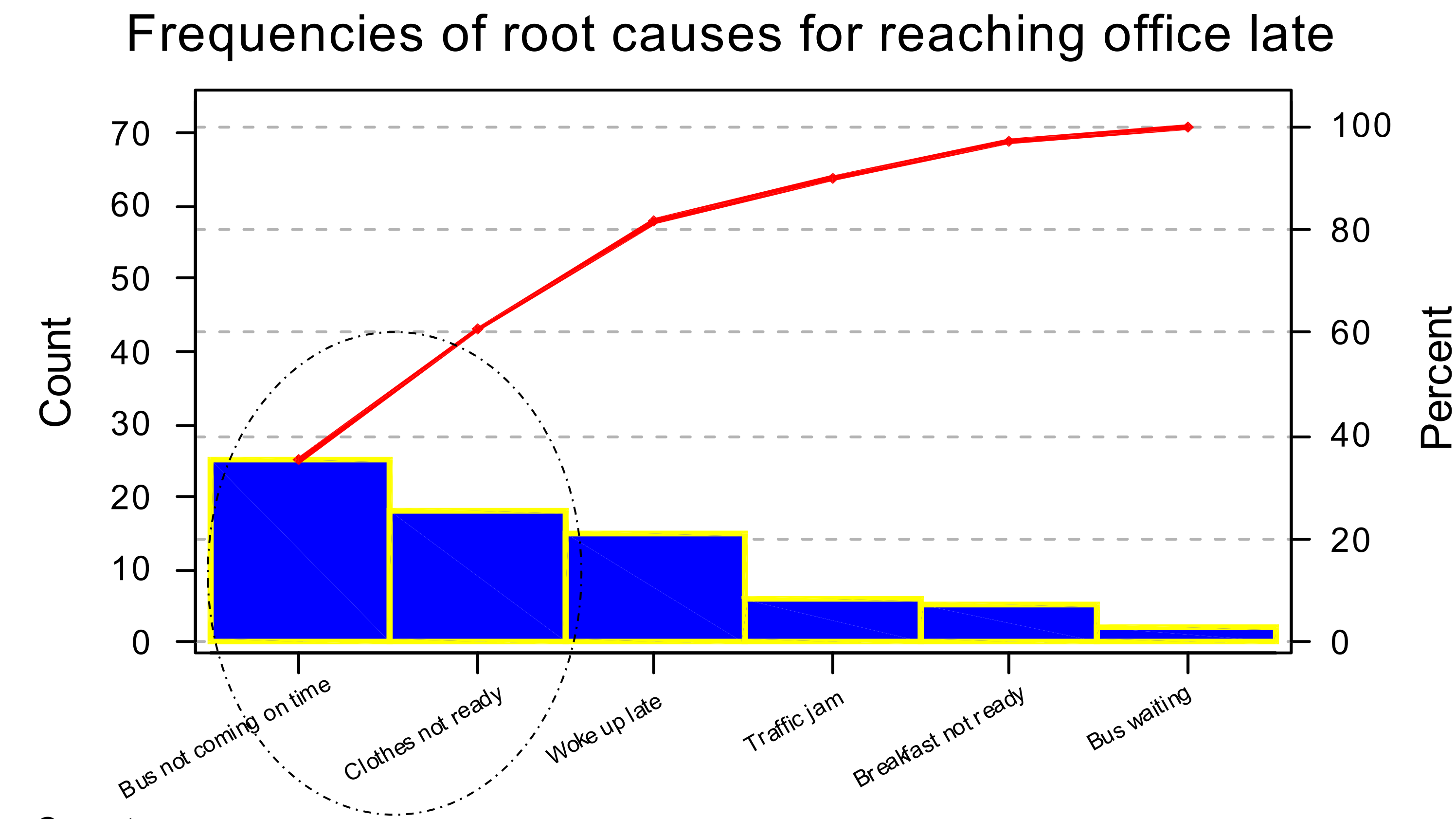
Pareto Chart



Pareto Diagram

- Suppose a person identifies multiple root-causes of reaching his office late. Now he is not sure where to focus so that he reduces the occurrence of reaching late by minimum 50%.
- He has identified following root causes
  - Woke up late
  - Clothes not ready
  - Breakfast not ready
  - Bus not coming on time
  - Traffic jam
  - Bus waiting for other employees
- He collects data on how frequent each of the root cause is & constructs a Pareto

Pareto Diagram



Count	25	18	15	6	5	2
Percent	35.2	25.4	21.1	8.5	7.0	2.8
Cum %	35.2	60.6	81.7	90.1	97.2	100.0

Pareto Diagram

- Essentially, Pareto is used to prioritize the problem areas / root-causes
- However, it can also be used to segment project defects to get clues about the process behavior
- Common factors used for segmentation are as below:

Factor	**Example**
What	Complaints, Defects, Problems
When	Year, Month, Week, Day
Where	Country, Region, City, Work Site
Who	Business, Department, Individual

Hypotheses Testing

Introduction

Always about a **population parameter**

Attempt to prove (or disprove) some assumption

Setup:

alternate hypothesis: What you wish to prove

Example: Person is guilty of crime

null hypothesis: Assume the opposite of what is to be proven. The null is always stated as an equality.

Example: Person is innocent

Hypothesis testing

What is Hypothesis Testing?

In our judicial system, a man is considered innocent until proven guilty

Based on the verdict, the following scenarios are possible

Truth	**Verdict**	
Innocent	Innocent	✓
Guilty	Guilty	✓
Innocent	Guilty	α or Type 1 error
Guilty	Innocent	β or Type 2 error

Which type of error is more serious?

Null Hypothesis

- When a person is being prosecuted for a crime, the judge hears the proceedings assuming that the person has committed no crime
- The job of the prosecutor is to prove his assumption wrong
- In other words, the person is non-guilty till proven otherwise, i.e. status quo
- Assuming status quo is *Null Hypothesis*

Alternative Hypothesis

- Alternative hypothesis challenges the null hypothesis
- If null hypothesis is proven wrong, alternative hypothesis must be right
- The prosecutor believes in the alternative hypothesis & gives proofs to substantiate it

Type I & Type II Errors

- Rejecting a null hypothesis when it was true is called Type I or ‘α’ error
  - It is also called ‘Producer’s Risk’ by drawing analogy with a part getting rejected by QA
  - team when it was not defective, thereby bringing loss to producer
  - Thus, concluding that coach B is better than coach A when they are actually at the same
  - level of efficiency, is making an ‘α’ error
- Accepting a null hypothesis when it was false is called Type II or ‘β’ error
  - It is also called ‘Consumer’s Risk’ by drawing analogy with a part getting accepted by QA
  - team when it was defective, thereby bringing loss to consumer who will buy that part
  - Thus, concluding that coach A & coach B are at the same level of efficiency when actually
  - they are not, is making a ‘β ’ error
- Probability of making one type of error can be reduced only when we are willing to accept a higher probability of making other type of error

  Usually, $\alpha = 0.05$ & $\beta = 0.10$

- http://www.mathnstats.com/index.php/hypothesis-testing/112-reject-or-fail-to-reject.html

Type I & Type II Errors

		Reality (Population)	
		Good	**Bad**
Evidence (Sample)	**Good**	**OK**	**Not OK** Type II Error Consumer's Risk
	Bad	**Not OK** Type I error Producer Risk	**OK**

Hypothesis Testing

Two complementary statements:

- H_0: Null Hypothesis.
- H_a: Alternative Hypothesis

Only one of them can be True

In the previous example, our Null Hypothesis is:

Ho: The defendant is Innocent

Ha: The defendant is Guilty

To avoid making a Type 1 error, the judge / jury needs to be 95% sure that the evidence points towards the man's guilt.

The P-Value Review

- Alpha is the probability of making a Type I error

- **The p-value is the probability of getting the observed difference or greater when Ho is true**

- Unless there is an exception based on engineering judgment, we will set an acceptance level of a Type I error at a = 0.05

- **<u>Thus, any p-value less than 0.05 means we reject the null hypothesis</u>**

An Example

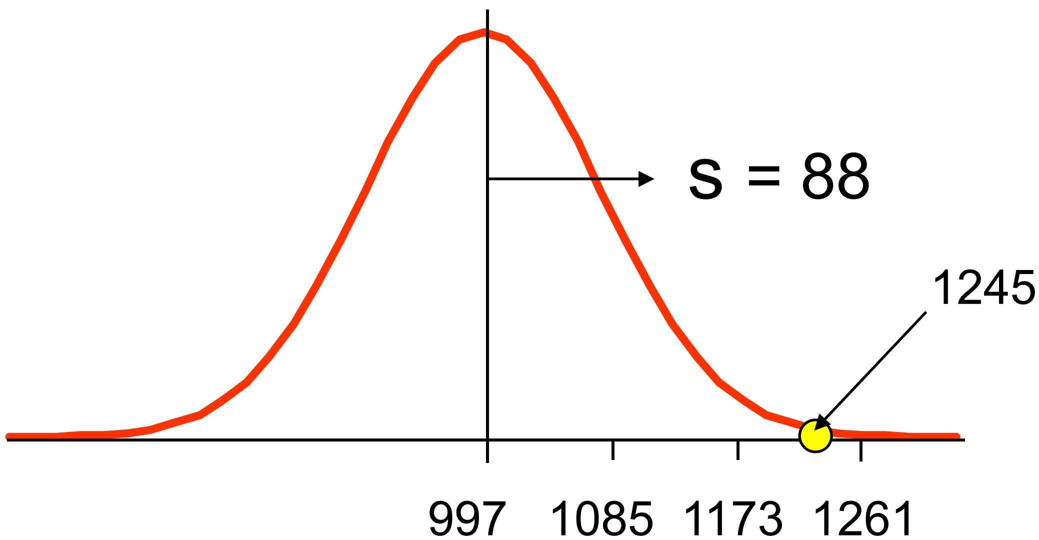
- You manage a warranty claims department. A customer claims loss of earnings of \$1,245 for an item which usually is about \$1,000.
- You examine 250 previous claims of the same item to make a comparison and find the average to indeed be \$997 with a standard deviation \$88.
- You want to know if the customer is over-claiming or if it is reasonable

Innocent or Guilty?

- If the customer is not over-claiming (it is a legitimate claim) then we would expect the claim to fit with the pattern of data represented by the previous 250 claims
- If the customer is over-claiming then we would expect the claim to not fit the pattern of data from the previous 250 claims

Does the Claim fit the Pattern?

You remember from previous module that if the data is normally distributed you can apply normal theory



By using standard Normal tables we can calculate the probability of a claim of \$1245 based on the historical data. If the probability is low, then we can assume an Illegitimate claim

$$Z = \frac{X - X}{s}$$

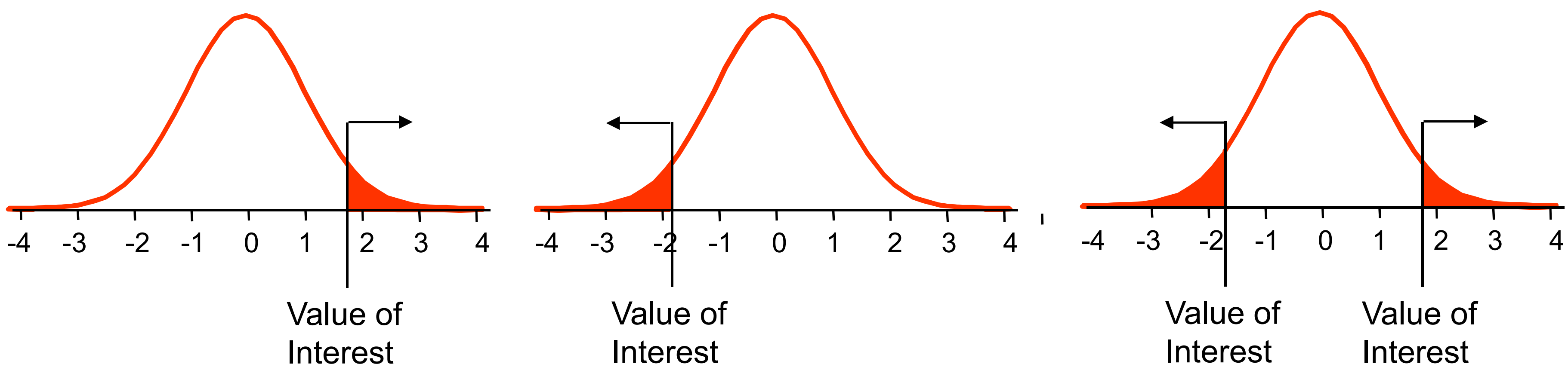
$$Z = \frac{1245 - 997}{88} = 2.82$$

From Standard Normal tables the P-value, the probability of being equal to or greater than 1245 is 0.0024 or 0.24%. In other words we would expect such a claim to happen 1 in 417 claims

P-values are Probabilities of Interest

P-value

- Tail area
- Area under curve beyond value of interest
- Probability of being at value of interest or beyond



Making the Decision

- The Normal theory is telling us that based on the previous 250 claims, we should expect a claim of $1245 or greater every 417 claims Hence:
  - This claim is legitimate – it is that 1 in 417
  - This is not legitimate - it does not fit the previous data

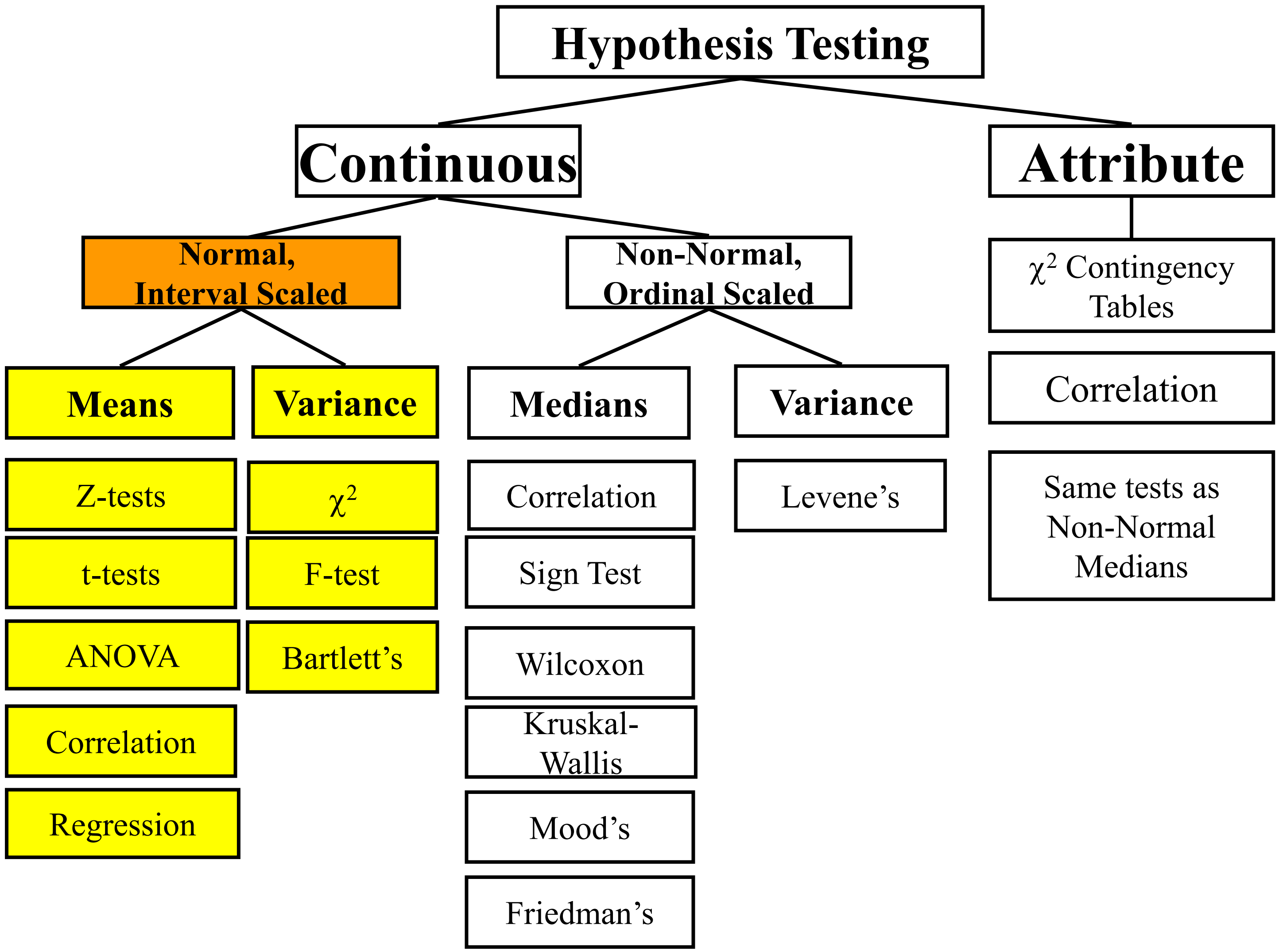
- Experience shows a small p-value (0 to 0.05) means
  - The probability is small that the value of interest comes from that distribution by chance therefore something else is going on
  - Since our p-value $0.024 < 0.05$ we can conclude that the claim is not legitimate

What Have we done?

- In the previous example we have used the properties of the Normal distribution to test whether the occurrence of an event could have happened by chance (the data fits the expected pattern) or there is a real difference (the data does not fit the expected pattern)

- This type of situation occurs frequently during Six Sigma improvement projects, either in the
  - Analyze phase when we are looking for differences to identify potential roots causes
  - Improve and Control phases when we are aiming to demonstrate that a real change has been made – we have made a difference

Hypothesis Testing Roadmap



Parametric Tests

- Use parametric tests when:

- The data are normally distributed
- The variances of populations (if more than one is sampled from) are equal
- The data are at least interval scaled

One sample z - test

Used when testing to see if sample comes from a known population. A sample of 25 measurements shows a mean of 17. Test whether this is significantly different from a the hypothesized mean of 15, assuming the population standard deviation is known to be 4.

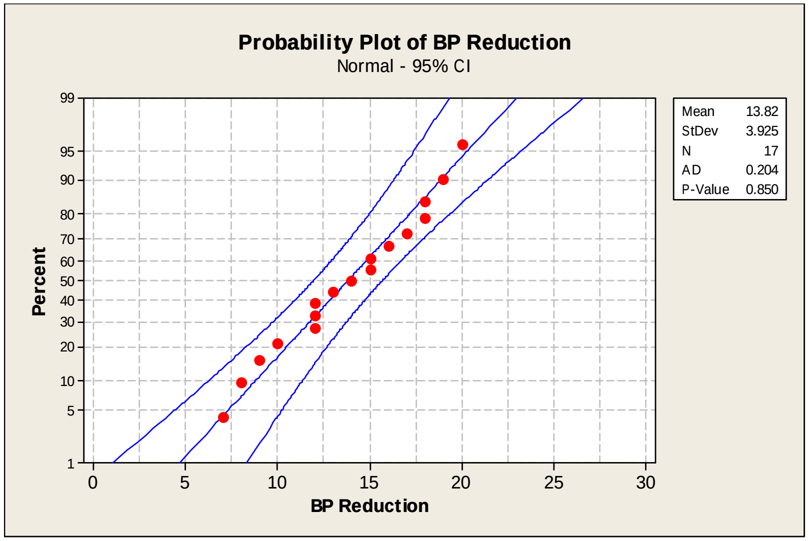
One-Sample Z

Test of mu = 15 vs not = 15
The assumed standard deviation = 4

N	Mean	SE Mean	95% CI	Z	P
25	17.0000	0.8000	(15.4320, 18.5680)	2.50	0.012

One sample t-test

BP Reduction%
10
12
9
8
7
12
14
13
15
16
18
12
18
19
20
17
15



The data show reductions in Blood Pressure in a sample of 17 people after a certain treatment. We wish to test whether the average reduction in BP was at least 13%, a benchmark set by some other treatment that we wish to match or better.

One Sample t-test – Minitab results

One-Sample T: BP Reduction

Test of mu = 13 vs > 13

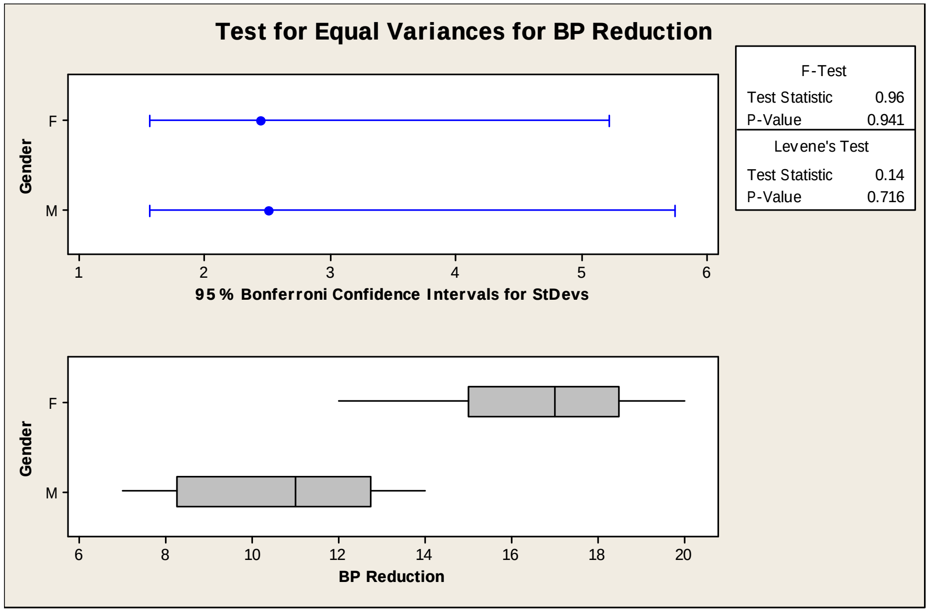
Variable	N	Mean	StDev	SE Mean	95% Lower Bound	T	P
BP Reduction	17	13.8235	3.9248	0.9519	12.1616	0.87	0.200

The p-value of 0.20 indicates that the reduction in BP could not be proven to be greater than 13%. There is a 0.20 probability that it is not greater than 13%.

Two Sample t-test

You realize that though the overall reduction is not proven to be more than 13%, there seems to be a difference between how men and women react to the treatment. You separate the 17 observations by gender, and wish to test whether there is in fact a significant difference between genders.

M	F
10	15
12	16
9	18
8	12
7	18
12	19
14	20
13	17
	15



Two Sample t-test

The test for equal variances shows that they are not different for the 2 samples. Thus a 2-sample t test may be conducted. The results are shown below. The p-value indicates there is a significant difference between the genders in their reaction to the treatment.

Two-sample T for BP Reduction M vs BP Reduction F

	N	Mean	StDev	SE Mean
BP Red M	8	10.63	2.50	0.89
BP Red F	9	16.67	2.45	0.82

Difference = mu (BP Red M) - mu (BP Red F)
Estimate for difference: -6.04167
95% CI for difference: (-8.60489, -3.47844)
T-Test of difference = 0 (vs not =): T-Value = -5.02 **P-Value = 0.000** DF = 15
Both use Pooled StDev = 2.4749

Paired t-test

- A paired t-test is not really like a 2-sample t-Test at all
- It is only a 1-sample t-Test in disguise
- The roadmap here is exactly the same as the 1-sample t-Test but applied instead to the differences between each pairing in the sample data sets
  - The target value in this case is zero

Do males earn higher average starting salaries than females?

(in $1,000s)	Males	Females
	60	32
	32	44
	80	22
	50	40
Sample Average:	$55.5	$34.5

Real question is whether males and females <u>in the same job</u> earn different average salaries.

It would be better to compare the difference in salaries in "**pairs**" of males and females.

Now, a Paired Study

Salaries (in $1,000s)

Job	Males	Females	Difference=M-F
Non-Profit	22	20	2.0
Education	29	28	1.0
Doctor	80	78	2.0
Scientist	35	32	3.0
Averages	41.5	39.5	2.0

P-value = How likely is it that a paired sample would have a difference as large as $2,000 if the true difference were 0?

Problem reduces to a One-Sample T-test on differences!!!!

Hypotheses Testing for Multiple Sample means

One-way Analysis of Variance (ANOVA)

ANOVA

- This method was developed by Sir Ronald Fisher in the 1930s as a way to interpret the results from agricultural experiments

- ANOVA is a statistically based, objective decision making tool for detecting any difference in the average performance of the groups of items tested

- Analysis of variance is a mathematical technique in which total variation is decomposed into its appropriate components.

One-Way ANOVA

- One-way Analysis of Variance (ANOVA) is used to compare the means of two or more samples against each other to determine whether it is likely that the sample could come from populations with the same means
- This is similar to a 2-sample t-Test except that three or more samples can be examined with ANOVA
- ANOVA can also be used to examine multiple Xs at the same time. In this section we focus on the One-Way ANOVA, which examines just one X
- For example, a Team might need to determine if three processors take the same amount of time to perform a task
    - A single X: Processor
    - With 3 levels: 3 Processors
- Levels are sometimes also called Treatments

What does ANOVA do?

- At its simplest ANOVA tests the following hypotheses:
  - H_0: The means of all the groups are equal.
  - H_a: Not all the means are equal
    - ✓ doesn't say how or which ones differ.
    - ✓ Can follow up with "multiple comparisons"
  - Note: we usually refer to the sub-populations as "groups" when doing ANOVA

- The conditions required to validate the use of the ANOVA method are:
  - The populations being sampled are normally distributed
  - The populations being sampled are homoscedastic
  - The observations are independent

Experimental Design

- The **sampling plan** or **experimental design** determines the way that a sample is selected.
- In an **observational study,** the experimenter observes data that already exist. The **sampling plan** is a plan for collecting this data.
- In a **designed experiment,** the experimenter imposes one or more experimental conditions on the experimental units and records the response.

Definitions

- An **experimental unit** is the object on which a measurement or measurements) is taken.
- A **factor** is an independent variable whose values are controlled and varied by the experimenter.
- A **level** is the intensity setting of a factor.
- A **treatment** is a specific combination of factor levels.
- The **response** is the variable being measured by the experimenter.

The Analysis of Variance (ANOVA)

- All measurements exhibit **variability.**
- The total variation in the response measurements is broken into portions that can be attributed to various **factors**.
- These portions are used to judge the effect of the various factors on the experimental response.

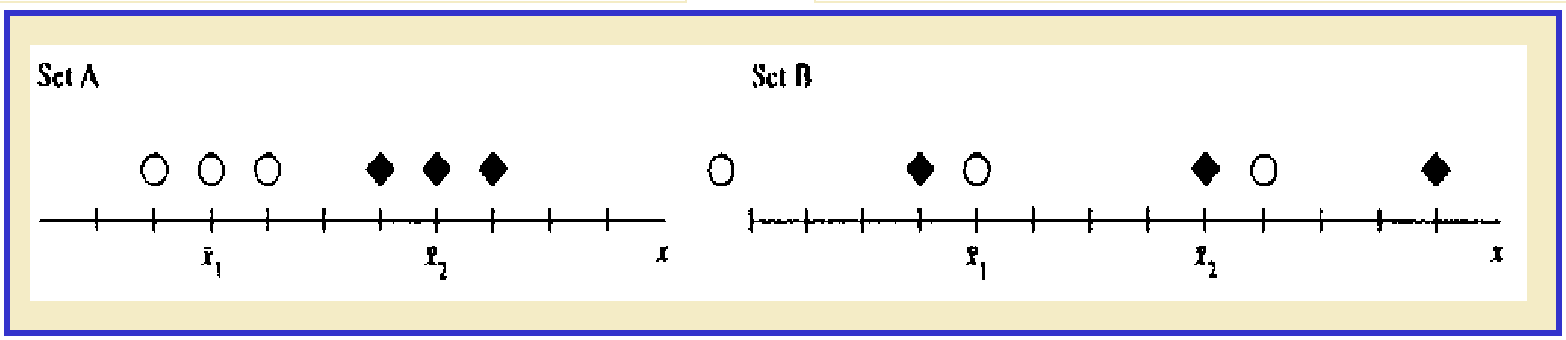
The Analysis of Variance

- If an experiment has been properly designed,

Total variation → **Factor 1**, **Factor 2**, **Random variation**

•We compare the variation due to any one factor to the typical random variation in the experiment.

The variation between the sample means is larger than the typical variation within the samples.	The variation between the sample means is about the same as the typical variation within the samples.



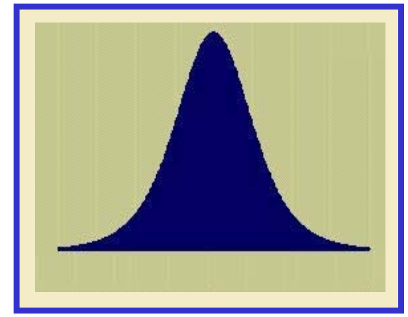
Assumptions

- Similar to the assumptions required in Chapter 10.

> - The observations within each population are normally distributed with a common variance s^{2}
> - 2. Assumptions regarding the sampling procedures are specified for each design.

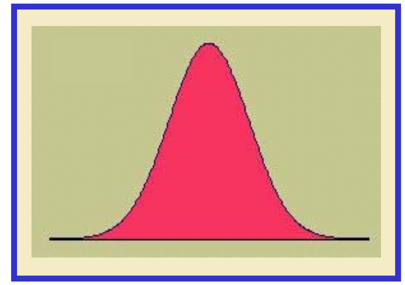
•Analysis of variance procedures are fairly robust when sample sizes are equal and when the data are fairly mound-shaped.

Three Designs

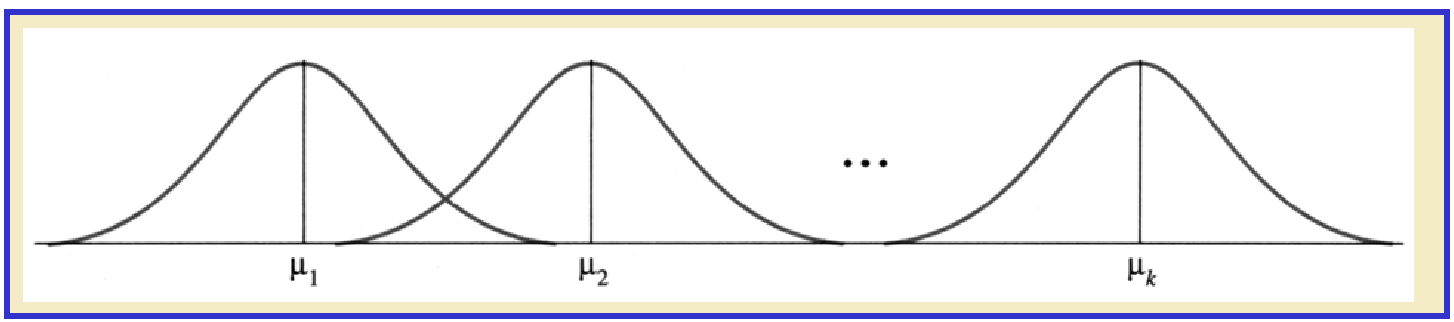


- **Completely randomized design:** an extension of the two independent sample *t*-test.
- **Randomized block design:** an extension of the paired difference test.
- ***a*** **×** ***b*** **Factorial experiment:** we study two experimental factors and their effect on the response.

The Completely Randomized Design



- A **one-way classification** in which one factor is set at k different levels.
- The k levels correspond to k different normal populations, which are the **treatments**.
- Are the k population means the same, or is at least one mean different from the others?



Example

Is the attention span of children affected by whether or not they had a good breakfast? Twelve children were randomly divided into three groups and assigned to a different meal plan. The response was attention span in minutes during the morning reading time.

No Breakfast	Light Breakfast	Full Breakfast
8	14	10
7	16	12
9	12	16
13	17	15

k = 3 treatments. Are the average attention spans different?

The Completely Randomized Design

- Random samples of size n_1, n_2, …,n_k are drawn from k populations with means m_1, m_2,…, m_k and with common variance s^2.
- Let x_{ij} be the j-th measurement in the i-th sample.
- The total variation in the experiment is measured by the **total sum of squares**:

$$\text{Total SS} = \sum (x_{ij} - \bar{x})^2$$

The Analysis of Variance

The **Total SS** is divided into two parts:

- **SST** (sum of squares for treatments): measures the variation among the *k* sample means.
- **SSE** (sum of squares for error): measures the variation within the *k* samples

in such a way that:

Computing Formulas

$$\text{Total SS} = \Sigma x_{ij}^2 - \text{CM}$$
$$= (\text{Sum of squares of all } x\text{-values}) - \text{CM}$$

with

$$\text{CM} = \frac{(\Sigma x_{ij})^2}{n} = \frac{G^2}{n}$$

$$\text{SST} = \Sigma \frac{T_i^2}{n_i} - \text{CM}$$

$$\text{SSE} = \text{Total SS} - \text{SST}$$

and

G = Grand total of all n observations

T_i = Total of all observations in sample i

n_i = Number of observations in sample i

$n = n_1 + n_2 + \cdots + n_k$

The Breakfast Problem

No Breakfast	Light Breakfast	Full Breakfast
8	14	10
7	16	12
9	12	16
13	17	15
$T_1 = 37$	$T_2 = 59$	$T_3 = 53$

G = 149

$$CM = \frac{149^2}{12} = 1850.0833$$

$$\text{Total SS} = 8^2 + 7^2 + \ldots + 15^2 - CM = 1973 - 1850.0833 = 122.9167$$

$$SST = \frac{37^2}{4} + \frac{53^2}{4} + \frac{59^2}{4} - CM = 1914.75 - CM = 64.6667$$

$$SSE = \text{Total SS} - SST = 58.25$$

Degrees of Freedom and Mean Squares

- These **sums of squares** behave like the numerator of a sample variance. When divided by the appropriate **degrees of freedom**, each provides a **mean square**, an estimate of variation in the experiment.
- **Degrees of freedom** are additive, just like the sums of squares.

$$\text{Total } df = \text{Trt } df + \text{Error } df$$

The ANOVA Table

Total *df* = $n_1+n_2+\ldots+n_k - 1 = n - 1$

Treatment *df* = $k - 1$

Error *df* = $n - 1 - (k - 1) = n - k$

Mean Squares

MST = SST/(k-1)

MSE = SSE/(n-k)

Source	df	SS	MS	F
Treatments	k -1	SST	SST/(k-1)	MST/MSE
Error	n - k	SSE	SSE/(n-k)	
Total	n -1	Total SS		

The Breakfast Problem

$$CM = \frac{149^2}{12} = 1850.0833$$

$$\text{Total SS} = 8^2 + 7^2 + \ldots + 15^2 - CM = 1973 - 1850.0833 = 122.9167$$

$$SST = \frac{37^2}{4} + \frac{53^2}{4} + \frac{59^2}{4} - CM = 1914.75 - CM = 64.6667$$

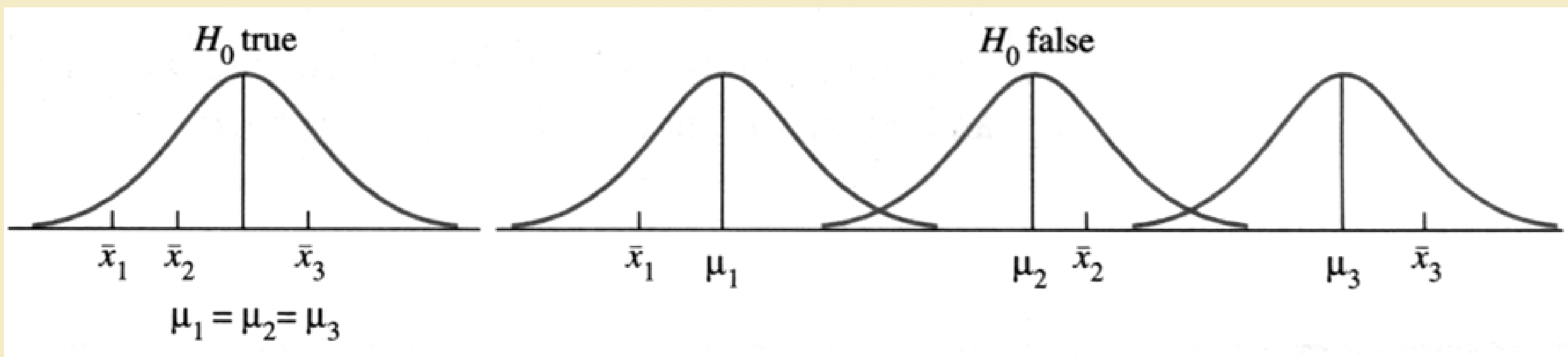
$$SSE = \text{Total SS} - SST = 58.25$$

Source	df	SS	MS	F
Treatments	2	64.6667	32.3333	5.00
Error	9	58.25	6.4722	
Total	11	122.9167		

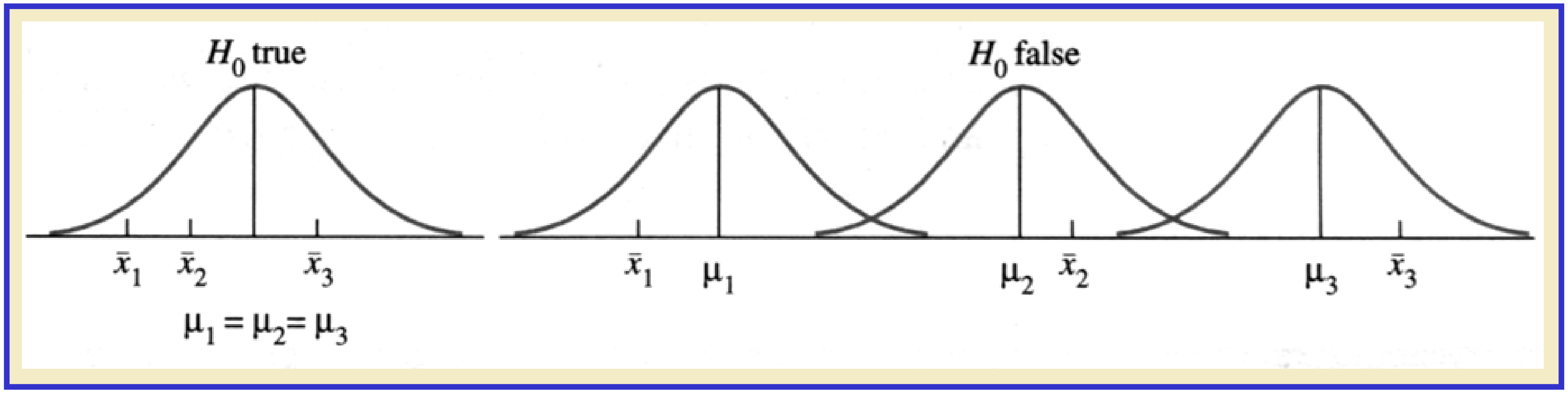
Testing the Treatment Means

$$H_0: \mu_1 = \mu_2 = \mu_3 = \ldots = \mu_k \text{ versus}$$

$$H_a: \text{at least one mean is different}$$



Remember that s^2 is the common variance for all k populations. The quantity MSE = SSE/(n - k) is a pooled estimate of s^2, a weighted average of all k sample variances, whether or not H_0 is true.

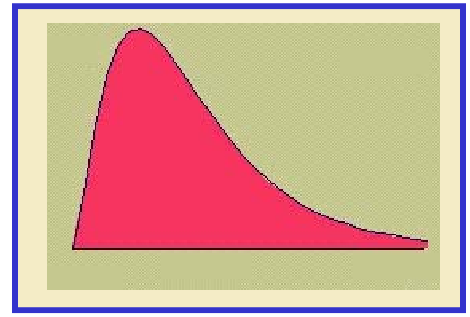


- If H_0 is true, then the variation in the sample means, measured by MST = [SST/ (k - 1)], also provides an unbiased estimate of s^2

- However, if H_0 is false and the population means are different, then MST— which measures the variance in the sample means — is unusually **large.** The test statistic **F = MST/ MSE** tends to be larger that usual

MY APPLET

The F Test

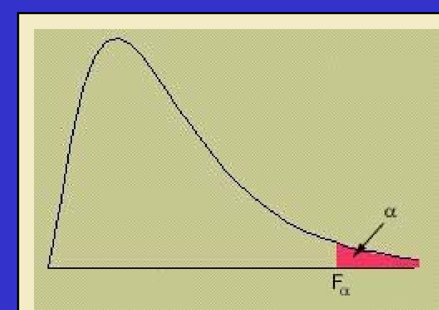


- Hence, you can reject H_0 for large values of F, using a **right-tailed** statistical test.

- When H_0 is true, this test statistic has an F distribution with $df_1 = (k - 1)$ and $df_2 = (n - k)$ degrees of freedom and **right-tailed** critical values of the F distribution can be used.

$$\text{Test Statistic: } F = \frac{MST}{MSE}$$

$$\text{Reject } H_0 \text{ if } F > F_\alpha \text{ with } k-1 \text{ and } n\text{-}k\ df.$$



The Breakfast Problem

Source	df	SS	MS	F
Treatments	2	64.6667	32.3333	5.00
Error	9	58.25	6.4722	
Total	11	122.9167		

MY APPLET

F-Table

F - Distribution ($\alpha = 0.05$ in the Right Tail)

Numerator Degrees of Freedom (df_1); Denominator Degrees of Freedom (df_2)

df_2 \ df_1	1	2	3	4	5	6	7	8	9
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385
3	10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123
4	7.7086	9.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	6.9988
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767
8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563
19	4.3807	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3660
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419
23	4.2793	3.4221	3.0280	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002
25	4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821
26	4.2252	3.3690	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655
27	4.2100	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501
28	4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360
29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2783	2.2229
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107
40	4.0847	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.0970	2.0401
120	3.9201	3.0718	2.6802	2.4472	2.2899	2.1750	2.0868	2.0164	1.9588
∞	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799

F - Distribution ($\alpha = 0.05$ in the Right Tail)

Numerator Degrees of Freedom (columns df_1); Denominator Degrees of Freedom (rows df_2)

df_2 \ df_1	10	12	15	20	24	30	40	60	120	∞
1	241.88	243.91	245.95	248.01	249.05	250.10	251.14	252.20	253.25	254.31
2	19.396	19.413	19.429	19.446	19.454	19.462	19.471	19.479	19.487	19.496
3	8.7855	8.7446	8.7029	8.6602	8.6385	8.6166	8.5944	8.5720	8.5494	8.5264
4	5.9644	5.9117	5.8578	5.8025	5.7744	5.7459	5.7170	5.6877	5.6581	5.6281
5	4.7351	4.6777	4.6188	4.5581	4.5272	4.4957	4.4638	4.4314	4.3985	4.3650
6	4.0600	3.9999	3.9381	3.8742	3.8415	3.8082	3.7743	3.7398	3.7047	3.6689
7	3.6365	3.5747	3.5107	3.4445	3.4105	3.3758	3.3404	3.3043	3.2674	3.2298
8	3.3472	3.2839	3.2184	3.1503	3.1152	3.0794	3.0428	3.0053	2.9669	2.9276
9	3.1373	3.0729	3.0061	2.9365	2.9005	2.8637	2.8259	2.7872	2.7475	2.7067
10	2.9782	2.9130	2.8450	2.7740	2.7372	2.6996	2.6609	2.6211	2.5801	2.5379
11	2.8536	2.7876	2.7186	2.6464	2.6090	2.5705	2.5309	2.4901	2.4480	2.4045
12	2.7534	2.6866	2.6169	2.5436	2.5055	2.4663	2.4259	2.3842	2.3410	2.2962
13	2.6710	2.6037	2.5331	2.4589	2.4202	2.3803	2.3392	2.2966	2.2524	2.2064
14	2.6022	2.5342	2.4630	2.3879	2.3487	2.3082	2.2664	2.2229	2.1778	2.1307
15	2.5437	2.4753	2.4034	2.3275	2.2878	2.2468	2.2043	2.1601	2.1141	2.0658
16	2.4935	2.4247	2.3522	2.2756	2.2354	2.1938	2.1507	2.1058	2.0589	2.0096
17	2.4499	2.3807	2.3077	2.2304	2.1898	2.1477	2.1040	2.0584	2.0107	1.9604
18	2.4117	2.3421	2.2686	2.1906	2.1497	2.1071	2.0629	2.0166	1.9681	1.9168
19	2.3779	2.3080	2.2341	2.1555	2.1141	2.0712	2.0264	1.9795	1.9302	1.8780
20	2.3479	2.2776	2.2033	2.1242	2.0825	2.0391	1.9938	1.9464	1.8963	1.8432
21	2.3210	2.2504	2.1757	2.0960	2.0540	2.0102	1.9645	1.9165	1.8657	1.8117
22	2.2967	2.2258	2.1508	2.0707	2.0283	1.9842	1.9380	1.8894	1.8380	1.7831
23	2.2747	2.2036	2.1282	2.0476	2.0050	1.9605	1.9139	1.8648	1.8128	1.7570
24	2.2547	2.1834	2.1077	2.0267	1.9838	1.9390	1.8920	1.8424	1.7896	1.7330
25	2.2365	2.1649	2.0889	2.0075	1.9643	1.9192	1.8718	1.8217	1.7684	1.7110
26	2.2197	2.1479	2.0716	1.9898	1.9464	1.9010	1.8533	1.8027	1.7488	1.6906
27	2.2043	2.1323	2.0558	1.9736	1.9299	1.8842	1.8361	1.7851	1.7306	1.6717
28	2.1900	2.1179	2.0411	1.9586	1.9147	1.8687	1.8203	1.7689	1.7138	1.6541
29	2.1768	2.1045	2.0275	1.9446	1.9005	1.8543	1.8055	1.7537	1.6981	1.6376
30	2.1646	2.0921	2.0148	1.9317	1.8874	1.8409	1.7918	1.7396	1.6835	1.6223
40	2.0772	2.0035	1.9245	1.8389	1.7929	1.7444	1.6928	1.6373	1.5766	1.5089
60	1.9926	1.9174	1.8364	1.7480	1.7001	1.6491	1.5943	1.5343	1.4673	1.3893
120	1.9105	1.8337	1.7505	1.6587	1.6084	1.5543	1.4952	1.4290	1.3519	1.2539
∞	1.8307	1.7522	1.6664	1.5705	1.5173	1.4591	1.3940	1.3180	1.2214	1.0000

Hypotheses Testing Tools

Non-Parametric Tests

Parametric and Non Parametric

- **Parametric tests**
  - They are based on a model that involves a specific distribution (usually a normal distribution).
  - Hypothesis concern the parameters of this distribution such as the mean m or the variance s2.
  - The parameters are measured at least on an interval scale.

- **Non-parametric tests**
  - They do not make the same type of assumptions concerning the type of measurement or the specific form of the distribution.
  - The assumptions are very general.

- **Power**
  - The power of parametric tests generally is higher than for non-parametric tests. Parametric tests require fewer observations, less time.

Common Statistical Tests

Some common statistical tests based on the Normality assumption and non Parametric Equivalents

Table 2: Common Statistical Tests For Normal & Non Parametric Data		
Assumes Normality	**No Assumption Required**	
One sample Z test	One sample Sign	
One sample t test	One sample Wilcoxon	
Two sample t test	Mann - Whitney	
One way ANOVA	Kruskal - Wallis Moods Median	
Randomized Block (Two way ANOVA Analysis)	The Friedman Test	

Hypotheses Testing Tools
Non-Parametric Tests

1-sample Sign test

One Sample Non-parametric: Sign Test

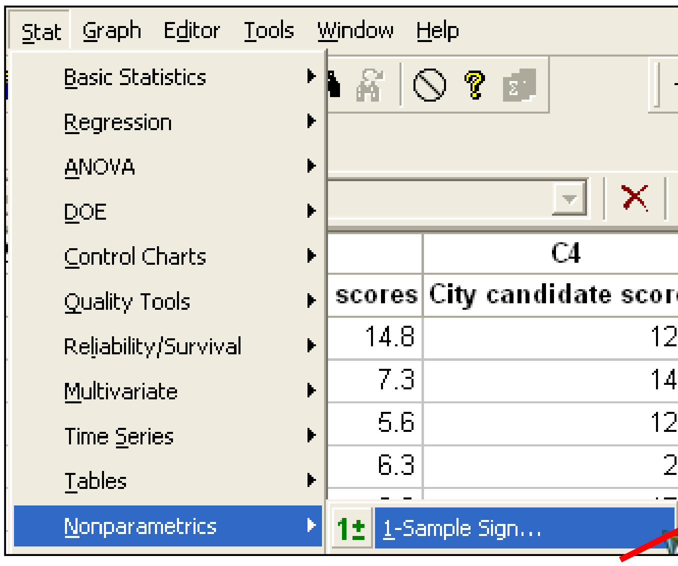
- **You can use the Sign test to perform a one sample sign test of the median or calculate the corresponding point estimate and confidence interval**

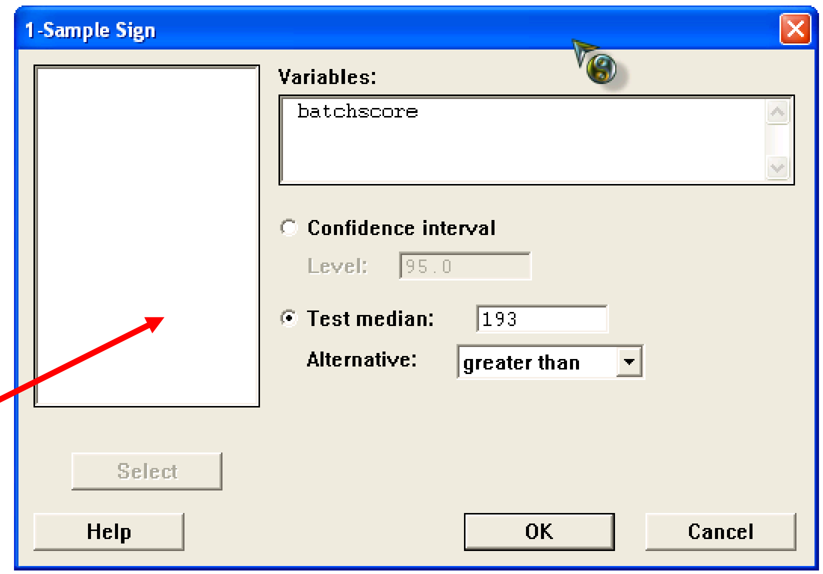
- **For the one-sample sign test, the hypotheses are**
  - H_0: median = hypothesized median versus
  - H_1: median $\neq$ hypothesized median

- **Use the sign test as a non-parametric alternative to one-sample Z-test and one-sample t-test which use mean rather than the median**

Minitab Exercise: Batch score data set: Sign Test

- As per a study students from training batches with median scores >193 have performed well on the job. Does the data confirm the hypothesis that the median score of the batch > 193?






```
Sign Test for Median: batchscore

Sign test of median =  193.0 versus > 193.0

              N  Below  Equal  Above       P  Median
batchscore   15      6      1      8  0.3953   195.0
```

Example

Price index values for 29 homes in a suburban area in the North were determined. Real estate records indicate the population median for similar homes the previous year was 115. This test will determine if there is sufficient evidence for judging if the median price index for the homes was greater than 115

Hypotheses Testing Tools
Non-Parametric Tests

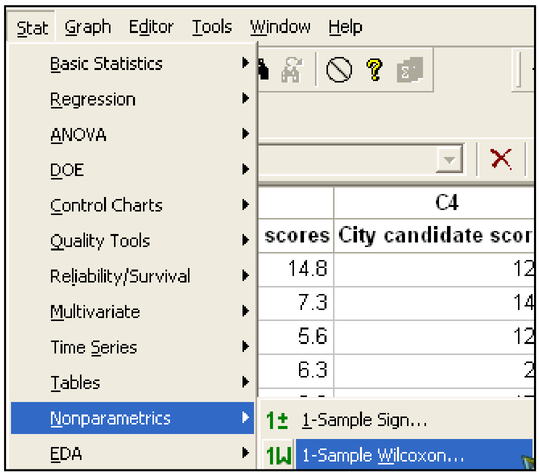
1-sample Wilcoxon test

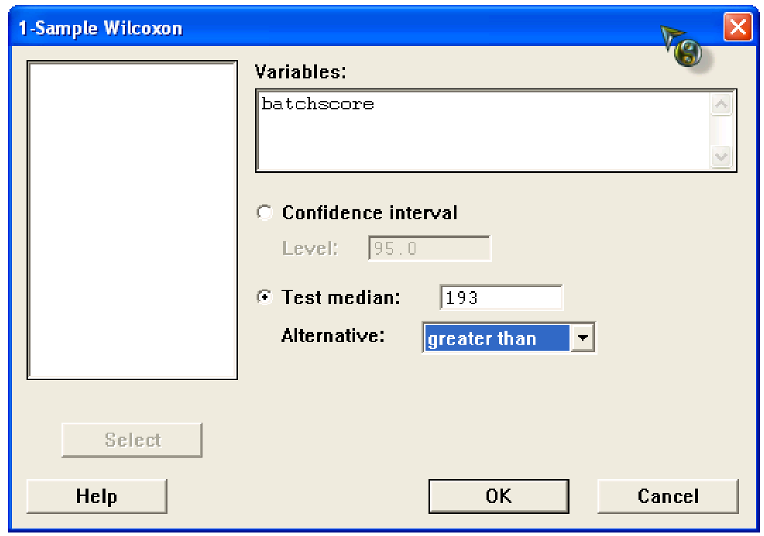
One Sample Non-parametric: Wilcoxon Test

- You can perform a one-sample Wilcoxon signed rank test of the median or calculate the corresponding point estimate and confidence interval.
- The Wilcoxon signed rank test hypotheses are
  - H_0: median = hypothesized median versus H_1: median $\neq$ hypothesized median
- **An assumption for the one-sample Wilcoxon test and confidence interval is that the data are a random sample from a continuous, symmetric population.**
- When the population is normally distributed, this test is slightly less powerful (the confidence interval is wider, on the average) than the t-test.
- It may be considerably more powerful (the confidence interval is narrower, on the average) for other populations.

Minitab Exercise: Batch score data set : Wilcoxon Test

- As per a study students from training batches with median scores >193 have performed well on the job. Does the data confirm the hypothesis that the median score of the batch > 193?





Wilcoxon Signed Rank Test: batchscore

Test of median = 193.0 versus median > 193.0

	N	N for Test	Wilcoxon Statistic	P	Estimated Median
batchscore	15	14	64.5	0.235	197.3

Example

Achievement test scores in science were recorded for 9 students. This test enables you to judge if there is sufficient evidence for the population median being different than 77 using a = 0.05.

Hypotheses Testing Tools
Non-Parametric Tests

2-sample Mann Whitney test

Two Sample Non-parametric: Mann Whitney Test

- You can perform a two-sample rank test (also called the Mann-Whitney test, or the two-sample Wilcoxon rank sum test) of the equality of two population medians, and calculate the corresponding point estimate and confidence interval.

- The hypotheses are: H0: $\eta 1 = \eta 2$ versus H1: $\eta 1 \neq \eta 2$ where η is the population median.

- **An assumption for the Mann-Whitney test is that the data are independent random samples from two populations that have the same shape (hence the same variance) and a scale that is continuous or ordinal (possesses natural ordering) if discrete.**

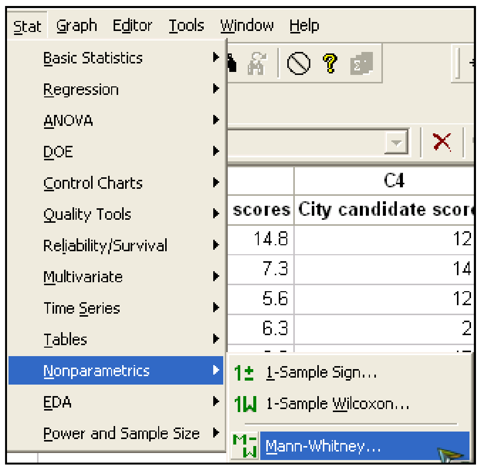
Two Sample Non-parametric: Mann Whitney Test

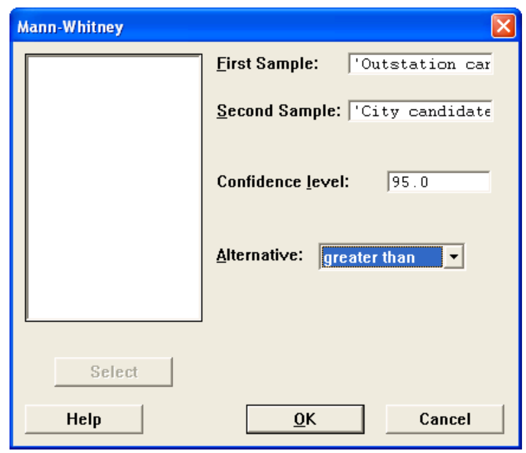
- The two-sample rank test is slightly less powerful (the confidence interval is wider on the average) than the two-sample test with pooled sample variance when the populations are normal, and considerably more powerful (confidence interval is narrower, on the average) for many other populations.

- If the populations have different shapes or different standard deviations, a two-sample t-test without pooling variances may be more appropriate.

Minitab Exercise: Outstation Batchscore data set : Mann Whitney Test (cont)

- A test was devised to see if outstation candidates performed better than those hired from within the city. Does the data confirm the hypothesis?





Mann-Whitney Test and CI: Outstation candidate scores, City candidate scores

```
                             N  Median
Outstation candidate scores  12   9.800
City candidate scores        36   7.750


Point estimate for ETA1-ETA2 is 0.800
95.1 Percent CI for ETA1-ETA2 is (-2.300,4.400)
W = 321.0
Test of ETA1 = ETA2 vs ETA1 > ETA2 is significant at 0.2640
The test is significant at 0.2639 (adjusted for ties)
```

Example

Samples were drawn from two populations and diastolic blood pressure was measured. You will want to determine if there is evidence of a difference in the population locations without assuming a parametric model for the distributions.

Hypotheses Testing Tools
Non-Parametric Tests

One-way Kruskal-Wallis test

One-Way Design Non-parametric: Kruskal-Wallis Test

- You can perform a Kruskal-Wallis test of the equality of medians for two or more populations.

- The Kruskal-Wallis hypotheses are:
  - H0: the population medians are all equal versus H1: the medians are not all equal

- An assumption for this test is that the samples from the different populations are independent random samples from continuous distributions, with the distributions having the same shape.

- **The Kruskal-Wallis test is more powerful than Mood's median test for data from many distributions, including data from the normal distribution, but is less robust against outliers.**

Minitab Exercise: InstructorScore data set: Kruskal-Wallis Test (cont)

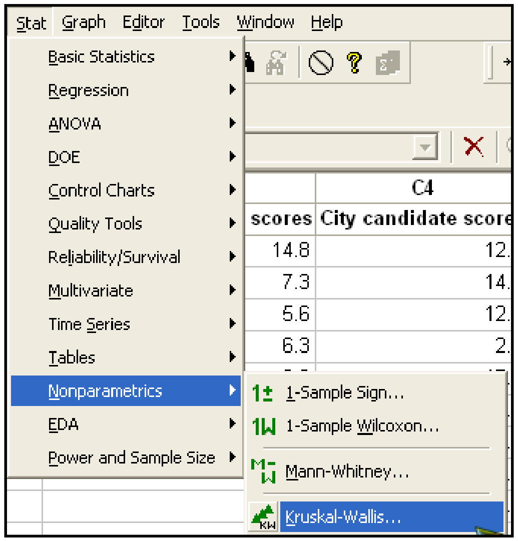
- Three instructors compared the grades they assigned over the past semester to see if some of them tended to give lower grades than others

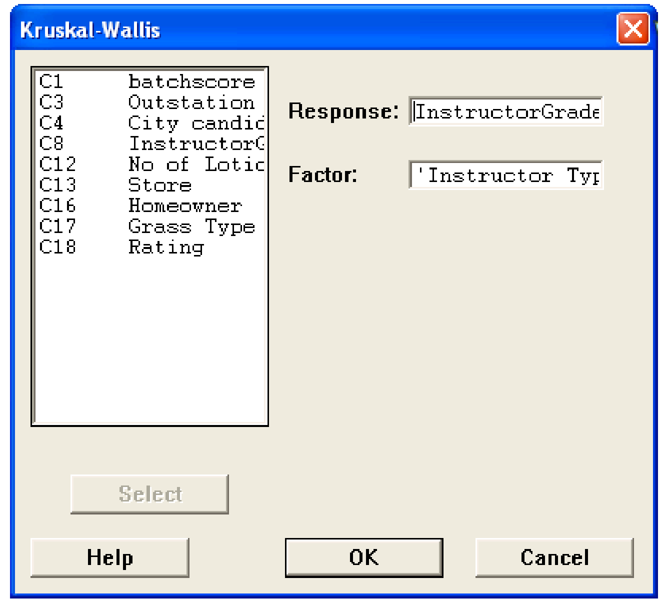
- The null hypothesis is: The three instructors grade evenly with each other

- The alternative of interest is: Some instructors tend to grade lower than others

- Does the data confirm the hypothesis?

Minitab Exercise: InstructorScore data set: Kruskal-Wallis Test (cont)





Kruskal-Wallis Test: InstructorGradeNo versus Instructor Type

```
Kruskal-Wallis Test on InstructorGradeNo

Instructor Type     N  Median  Ave Rank      Z
Instructor 1       43   3.000      54.9  -0.03
Instructor 2       38   3.000      53.3  -0.42
Instructor 3       28   3.000      57.6   0.50
Overall           109              55.0

H = 0.30  DF = 2  P = 0.860
H = 0.32  DF = 2  P = 0.852  (adjusted for ties)
```

Minitab Example

These are the scores of a sample of 20 student pilots on their Federal Aviation Agency written examination, The Mode of the examination are Video cassette, Audio Cassette and Classroom Training. The FAA is interested in evaluating the effectiveness of these training method.

Hypotheses Testing Tools
Non-Parametric Tests

One-way Mood’s Median test

One-Way Design Non-parametric: Mood's Median Test

- Mood's median test can be used to test the equality of medians from two or more populations.

- Mood's median test is sometimes called a median test or sign scores test.

- Mood's median test tests: H0: the population medians are all equal versus H1: the medians are not all equal

- An assumption of Mood's median test is that the data from each population are independent random samples and the population distributions have the same shape.

One-Way Design Non-parametric: Mood's Median Test

- Mood's median test is robust against outliers and errors in data and is particularly appropriate in the preliminary stages of analysis.

- Mood's median test is more robust than is the Kruskal-Wallis test against outliers, but is less powerful for data from many distributions, including the normal.

Minitab Exercise: Cartoon data set : Mood's Median Test (cont)

- One hundred seventy-nine participants were given a lecture with cartoons to illustrate the subject matter.

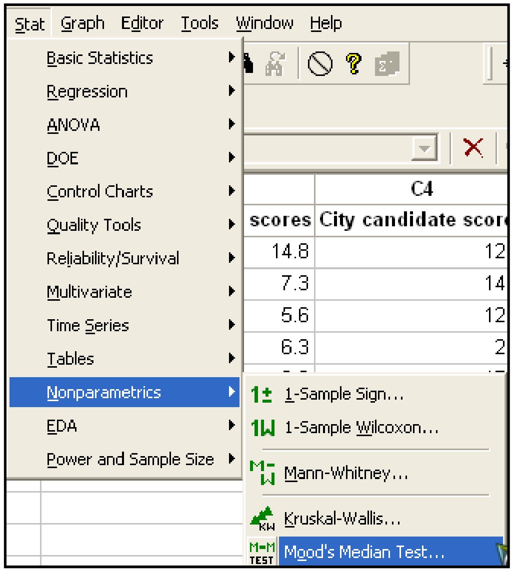
- Subsequently, they were given the OTIS test, which measures general intellectual ability.

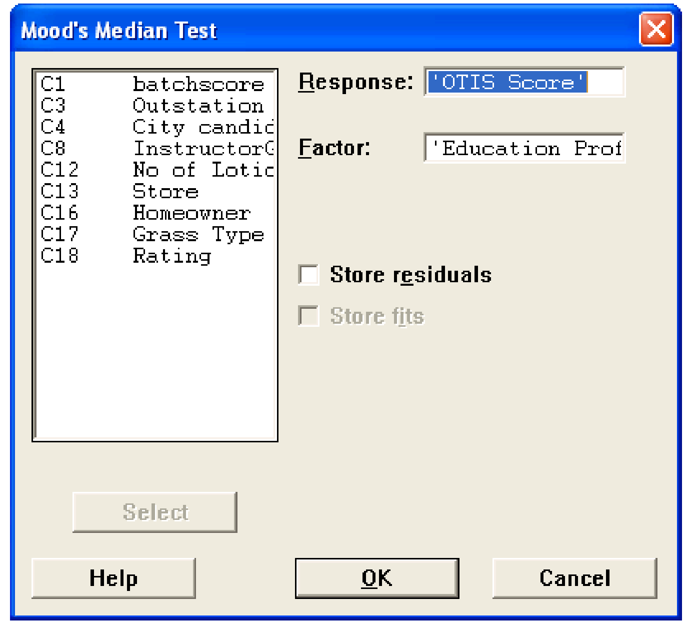
- Participants were rated by educational level:
  - 0 = preprofessional,
  - 1 =professional,
  - 2 = college student.

- The Mood's median test was selected to test H_0: $\eta_1 = \eta_2 = \eta_3$, versus H_1: not all η's are equal, where the η's are the median population OTIS scores for the three education levels.

- Does the data confirm the hypothesis?

Minitab Exercise: Cartoon data set : Mood's Median Test (cont)





Mood Median Test: OTIS Score versus Education Profile

```
Mood median test for OTIS Score
Chi-Square = 49.08    DF = 2    P = 0.000

Education                                Individual 95.0% CIs
Profile     N<=  N>  Median  Q3-Q1   ----+---------+---------+---------+--
0            47   9    97.5   17.3   (-----*-----)
1            29  24   106.0   21.5            (------*------)
2            15  55   116.5   16.3                            (----*----)
                                     ----+---------+---------+---------+--
                                      96.0     104.0     112.0     120.0

Overall median = 107.0
```

Friedman test

- Friedman test is a nonparametric analysis of a randomized block experiment, and thus provides an alternative to a Two-way analysis of variance. The hypotheses are:
- H0: all treatment effects are zero versus H1: not all treatment effects are zero
- Randomized block experiments are a generalization of paired experiments, and the Friedman test is a generalization of the paired sign test. Additivity (fit is sum of treatment and block effect) is not required for the test, but is required for the estimate of the treatment effects.

Example

A randomized block experiment was conducted to evaluate the effect of a drug treatment on enzyme activity. Three different drug therapies were given to four animals, with each animal belonging to a different litter.

Hypotheses Testing Discrete data

Chi-square analysis

Discrete X and Y Data

Discrete Ys

- Examples:
  - invoice accuracy (accurate, not accurate)
  - customer satisfaction (poor, fair, good, excellent)
  - types of application errors (wrong address, misspelled name, missing age, etc.)
- An attribute is recorded for each unit
- You can count the number of units with each attribute
- The counts are usually summarized in a table (known as a contingency table)

Discrete X and Y Data, contd.

Discrete Xs

- Examples
  - location (NY, LA, Denver)
  - method (Std, New)
  - product type (A, B, C, D)

- Separates the data into groups (it’s the stratifying or “by” variable)
  - “How do results vary by location?”

Chi-Square Test: What is it?

There is a statistical test designed to compare multiple proportions simultaneously Called Chi-Square test

$$\chi^2 = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$$

This test can be used to check for independence and homogeneity. They generally arise by one oft two sampling schemes

- We interview n people and classify them according to their responses to questions, i.e., one population
- We interview n_1 persons from one sub-population, n_2 persons from a separate sub-population, etc. and classify them according to their responses to a single question, i.e., several populations

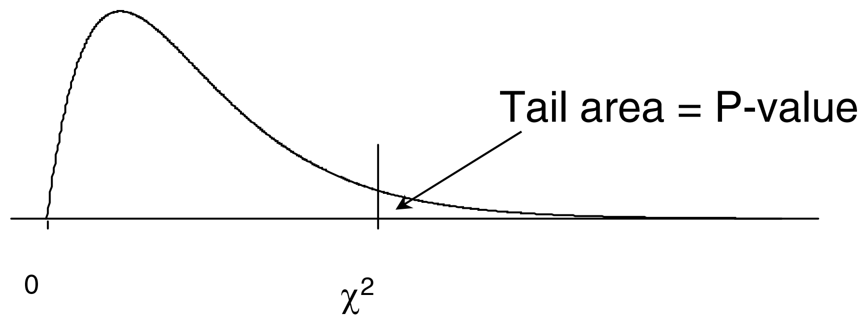
Chi-Square Test: What is it?, cont.

Case I is conducive to a test of independence (it allows the comparison of two attributes in a sample of data to determine if there is any relationship between them).

Case II leads to a test of homogeneity (it tests whether the proportions for each class are equal across all populations, i.e., the distribution of probabilities is homogeneous from population to population.)

Chi-Square Analysis

Chi-Square Distribution



- Skewed with a tail to the right
- Lower bound = 0
- Shape depends on degrees of freedom (df)
  - df = (# rows [r] – 1) x (# columns [c] – 1)
  - For our example: (2 – 1) x (4 – 1) = 3
- Provides the P-value
  - The probability that the difference between the observed and expected counts is due to random variation
- The larger the chi-square statistic, the smaller the P-value
  - Look at the distribution above: as c^2 gets larger, the tail area gets smaller

When to Use it?

Use of a Chi-square test is recommended on the following cases:

- To determine if the proportions among different groups from one population are similar

- To determine if there are differences in the response of different populations to a single question

How to Set Up a Test of Independence

Null Hypothesis (Ho):
The proportions of two or more groups are the same
Alternate Hypothesis (Ha):
At least one group proportion is different than the other group proportions

Decision criteria:

If the p-value is less than α (significance level), Ho is rejected.
At least one of the group proportions is significantly different than the others. Hence, there is independence
On the other hand, if the p-value is greater than α (significance level). This shows that there is not enough evidence to declare a statistically significant difference between the group
proportions

Assumptions of the Chi-Square Test

- The sample is representative of the population or process

- We assume the underlying distribution is binomial (or multinomial) for discrete data used in a χ^2 test

- The expected count ≥ 5 for each cell, or the test will not perform properly
  - If expected count is < 5, collecting additional data (a bigger sample size) may be needed

- Note: If expected cell count is less than 5, it is possible to use a modified χ^2 test, minimum acceptable expected value 2.

Example

The Personnel Department wants to see if there is a link between age (old and young) and whether that person gets hired.

What's the Y ? **Got Hired** Type of Data ? **Attribute**

What's the X ? **Age** Type of Data ? **Attribute**

What type of tool would you use ? **Chi-Square**

The Hypothesis

- With the Chi-Square Test for Independence, statisticians assume most variables in life are independent, therefore:
- H_0: Data is Independent (Not Related)
- (where Age & Hiring practices are independent)

  H_A: Data is Dependent (Related)
- (where Age & Hiring practices are dependent)

> The *p*-value is the probability that we are wrong in rejecting the null.

Hypothesis for Example

- Let's walk through our example ...
- Assume we wanted to determine if age and hiring practices are dependent or independent.
- Therefore our hypotheses are stated as follows ...
- H_0: Age and Hiring Practices are independent
- H_A: Age and Hiring Practices are dependent

Contingency Table

	Hired	Not Hired	Total
Old	30	150	180
Young	45	230	275
Total	75	380	455

Do Hiring Decisions depend on Age?

Step #1

- We must develop an Observed Frequency Table by breaking our attribute variables into different levels:
- Age: Old & Young
- Hiring Practices: Hired & Not Hired
- We then collect data to perform the analysis.

Example:	Hired	Not Hired
Old	30	150
Young	45	230

Step #2

Calculate Column & Row Totals

Example:	Hired	Not Hired	Total
Old	30	150	180
Young	45	230	275
Total	75	380	455

Step #3

- What would it look like if these factors were really independent?
- Develop an expected frequency table.

Example:

	Hired	**Not Hired**
Old		
Young		

How do we do that?

Step #3 Continued

- What would it look like if these factors were really independent?
- Develop an expected frequency table.

Example:	Hired	Not Hired	Total
Old	$\frac{75}{455} \times 180$		180
Young			275
Total	75	380	455

Cell's expected frequency is:

$$\frac{\text{(Column Total)} \times \text{(Row Total)}}{\text{Grand Total}}$$

Step #3 Continued

- We would expect the quantity of Old and Hired to be 29.6 if the two factors were really independent.

Example:	Hired	Not Hired	Total
Old	29.6	150.3	180
Young	45.3	229.7	275
Total	75	380	455

You finish the table!

Step #4

- Subtract the expected value from the observed (O-E)

Example:	Hired	Not Hired	Total
Old	30-29.6=0.4	-0.3	180
Young	-0.3	0.3	275
Total	75	380	455

You finish the table!

Step #5

- Square the Differences $(O-E)^2$

Example:	Hired	Not Hired	Total
Old	$(.4)\times(.4)=.16$	.09	180
Young	.09	.09	275
Total	75	380	455

You finish the table!

Step #6

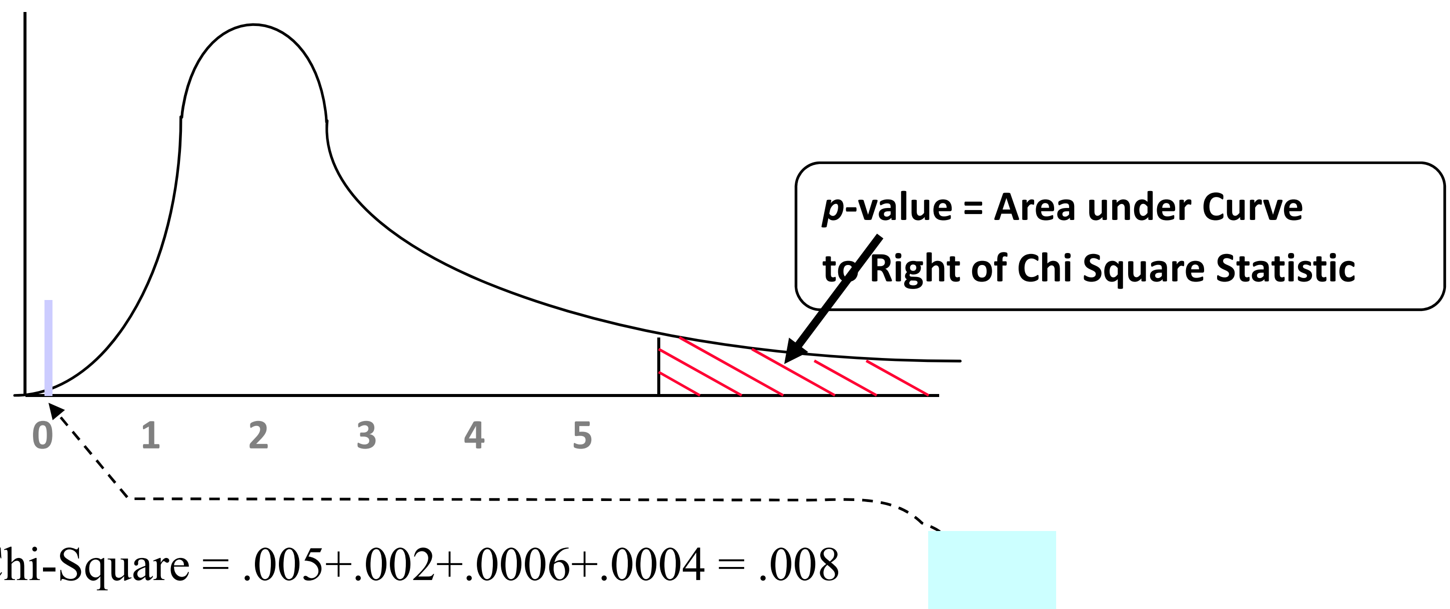
- Compute the Relative Squared Differences (O-E)^2 / E

Example:	Hired	Not Hired	Total
Old	.16 / 29.6 = .005	.0006	180
Young	.002	.0004	275
Total	75	380	455

You finish the table!

Chi-Square Distribution

- The sum of the relative squared differences is Chi-Square. Find Chi-Square on the distribution to determine significance.



Example: Chi-Square = .005+.002+.0006+.0004 = .008

Conclusion: Hiring Practices are Independent of Age

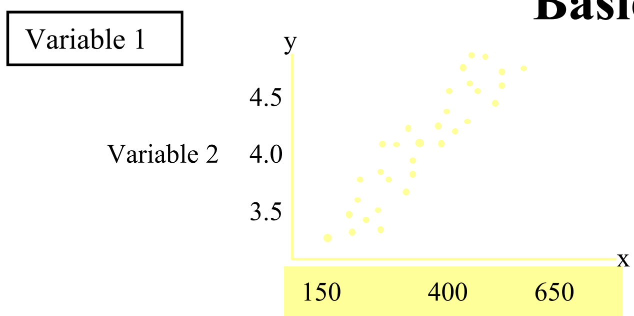
- If the 2 factors are independent, the sum of the differences will be close to 0
- The Larger the Chi-Square Statistic, the smaller the *p*-value, the more likely the variables are dependent.

Relationship of
$Y = f(x)$

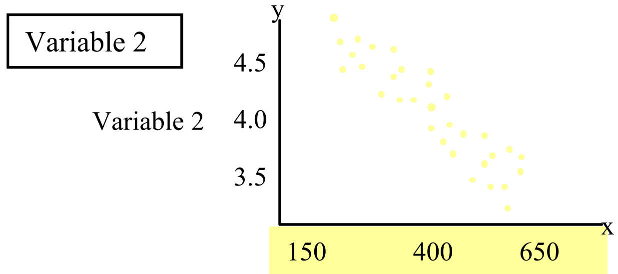
Scatter Plot
Correlation Analysis
Regression Analysis

SCATTER DIAGRAM

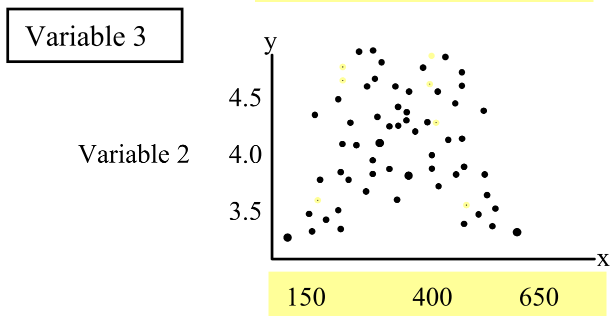
Basics of a Scatter Diagram



Positive Correlation
An increase in y may be
related to an increase in x.



Negative Correlation
A decrease in y may be
related to an increase in x.



No Correlation
There is no demonstrated
connection between y and x.

Definition: A Scatter Diagram is a tool that helps a team to identify and study the possible relationship between changes observed in two sets of variables, such as height and weight. Each point on the diagram represents a pair of measurements, one for each variable, plotted on X, Y coordinates.

ASQ India

Scatter Diagram Objectives

- To work as a team to test a hypothesis that two variables are related.
- To provide the team both graphical and statistical means to test the strength of the relationship between the two variables.
- To provide a data-rich follow-up to a Cause and Effect Diagram. To see if there is more than a consensus opinion that a suspected root cause really does influence the original problem.
- To gather data simultaneously on pairs of variables that seem to rise together, fall together, or move in opposite directions together.

ASQ India

Constructing the Scatter Diagram

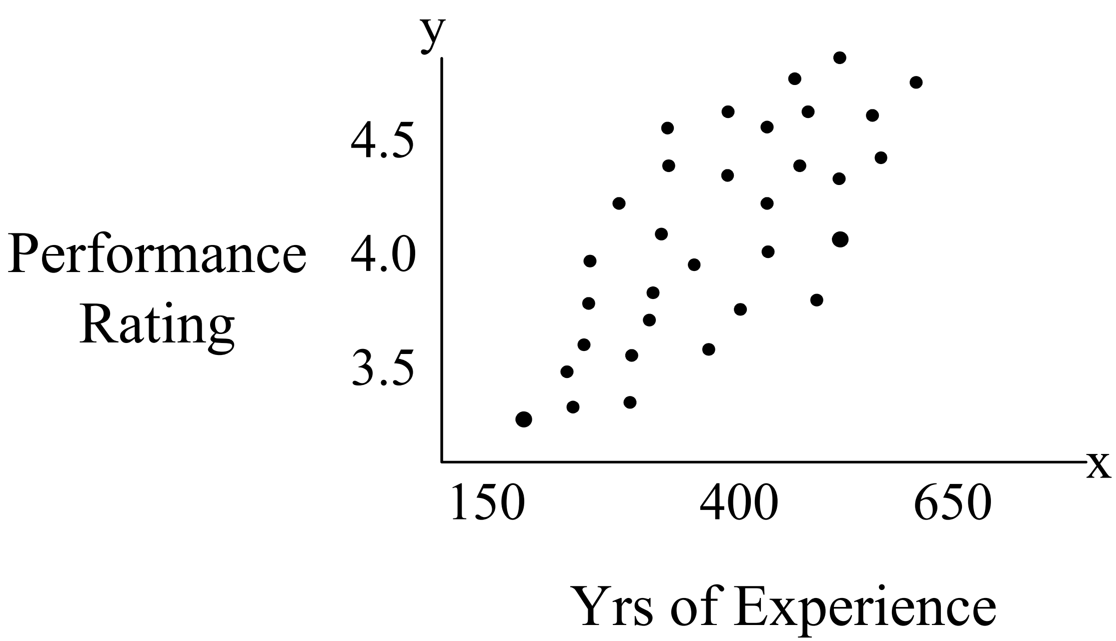
- Collect Paired Data
- Draw the Axes
- Plot the Data
- Interpret the Data

Exercise – Scatter Diagram

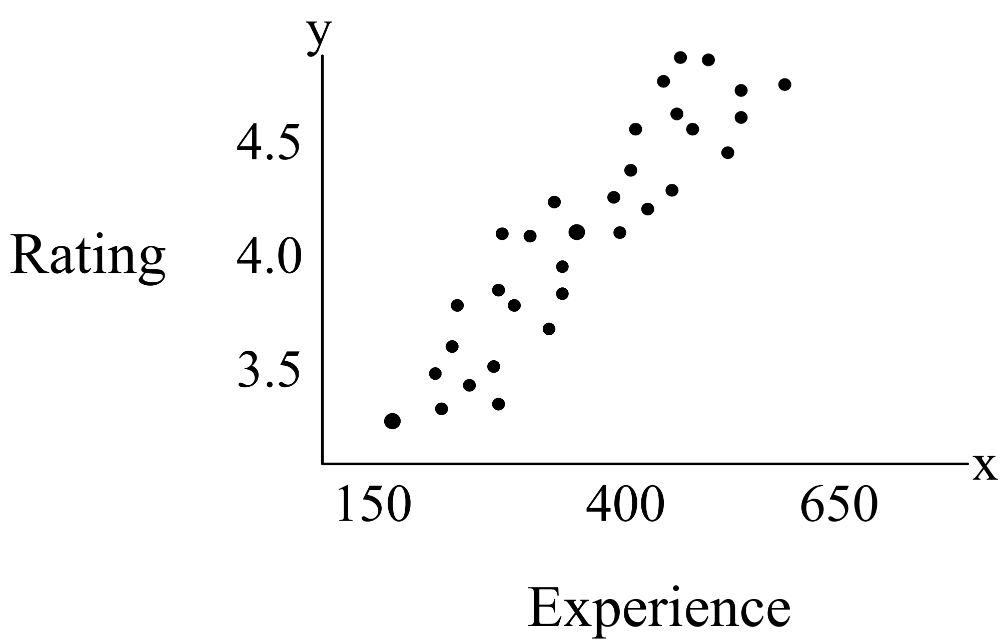
The marks received by participants in a competitive exam and number of hours coaching taken, are given below. Can you comment on this?

Hrs of Coaching	Marks Obtained
100	89
85	78
40	37
65	59
115	93
75	65

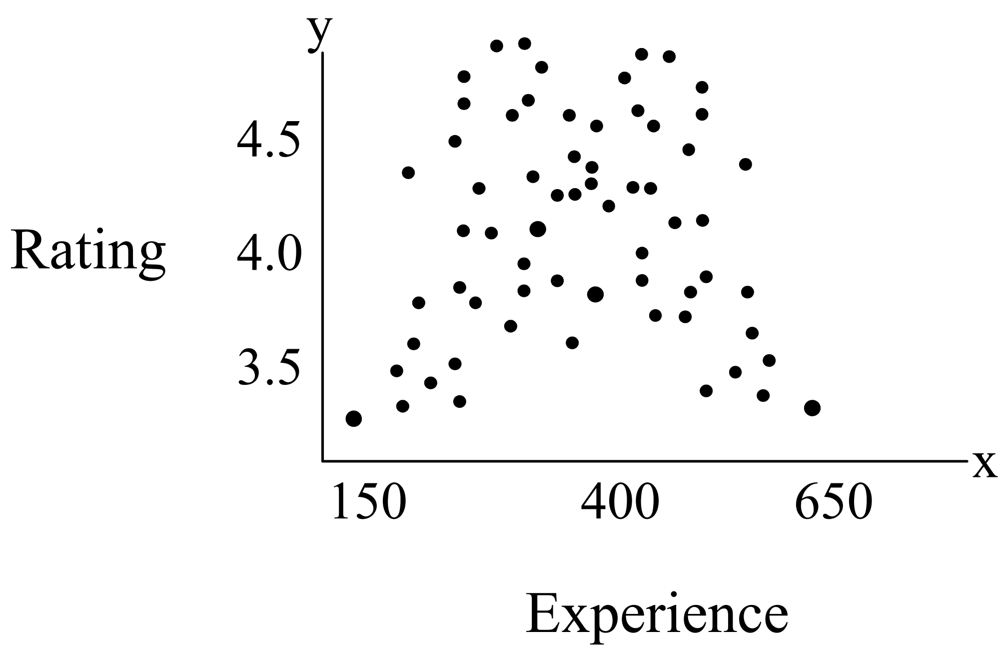
Possible Correlation



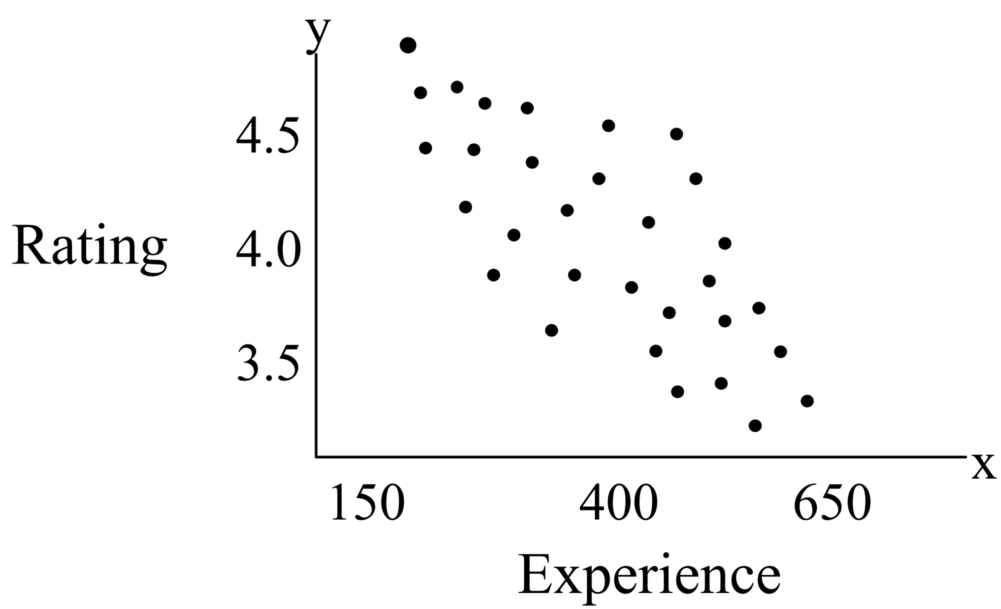
Positive Correlation



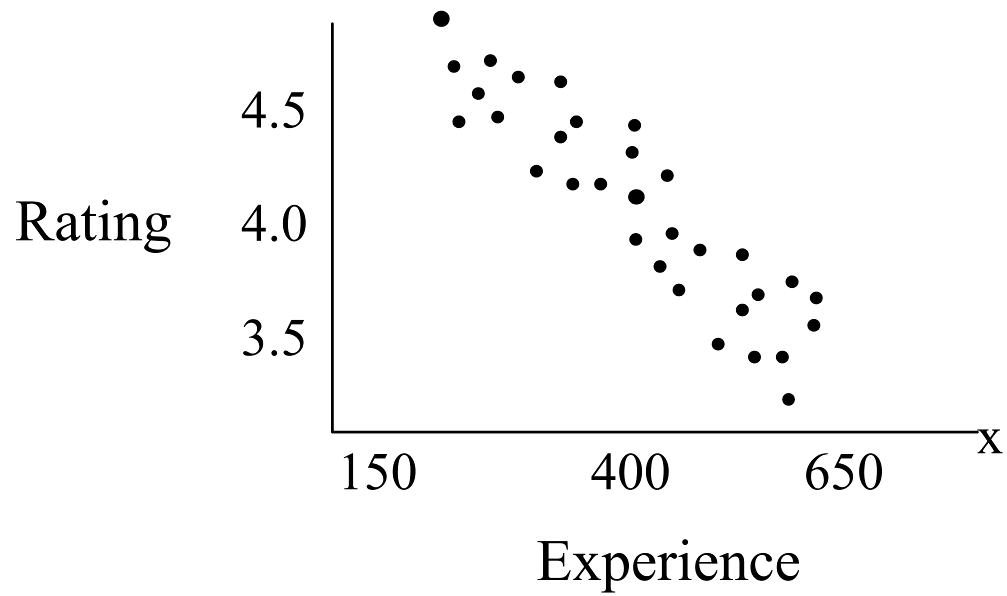
No Correlation



Possible Negative Correlation



Negative Correlation



The Correlation Coefficient

- The strength and direction of the relationship between *x* and *y* are measured using the **correlation coefficient, *r*.**

$$r = \frac{s_{xy}}{s_x s_y}$$

where

$$s_{xy} = \frac{\sum x_i y_i - \dfrac{(\sum x_i)(\sum y_i)}{n}}{n-1}$$

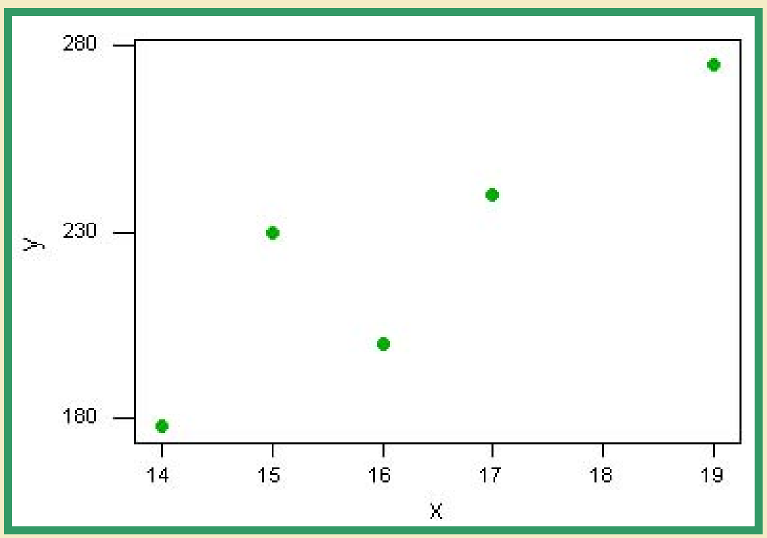
s_x = standard deviation of the *x*'s

s_y = standard deviation of the *y*'s

Example

- Living area *x* and selling price *y* of 5 homes.

Residence	1	2	3	4	5
x (thousand sq ft)	14	15	17	19	16
y ($000)	178	230	240	275	200



- The scatterplot indicates a positive linear relationship.

Example

x	y	xy
14	178	2492
15	230	3450
17	240	4080
19	275	5225
16	200	3200
81	1123	18447

Calculate

$\bar{x} = 16.2 \quad s_x = 1.924$

$\bar{y} = 224.6 \quad s_y = 37.360$

$$s_{xy} = \frac{\sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n}}{n-1}$$

$$= \frac{18447 - \frac{(81)(1123)}{5}}{4} = 63.6$$

$$r = \frac{s_{xy}}{s_x s_y}$$

$$= \frac{63.6}{1.924(37.36)} = .885$$

Interpreting ***r***

- $-1 \le r \le 1$ — Sign of r indicates direction of the linear relationship.

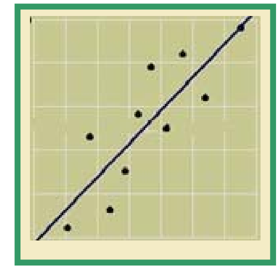
- $r \approx 0$ — Weak relationship; random scatter of points

- $r \approx 1$ or -1 — Strong relationship; either positive or negative

- $r = 1$ or -1 — All points fall exactly on a straight line.

Regression Analysis

A Simple Linear Model

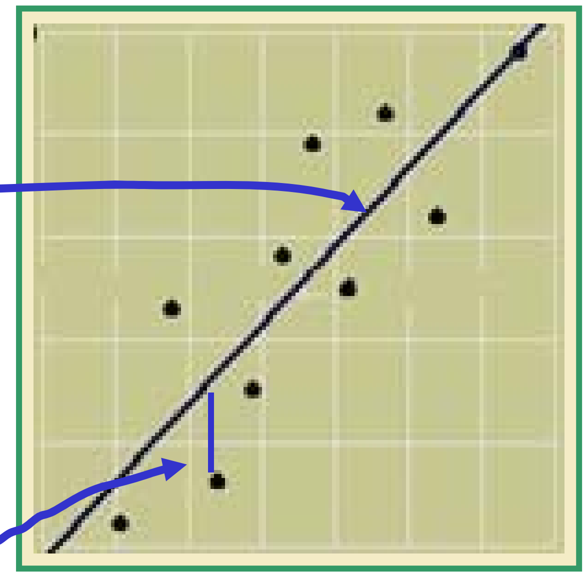


If we want to describe the relationship between y and x for the **whole population**, there are two models we can choose

- Deterministic Model: $y = a + bx$
- Probabilistic Model:
  - y = deterministic model + random error
  - $y = a + bx + e$

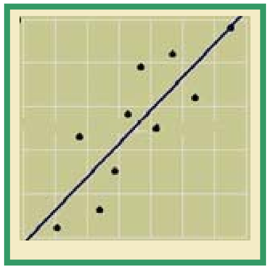
A Simple Linear Model

- Since the bivariate measurements that we observe do not generally fall **exactly** on a straight line, we choose to use:
- **Probabilistic Model:**
  - $y = \mathbf{a} + \mathbf{b}x + \mathbf{e}$
  - $E(y) = \mathbf{a} + \mathbf{b}x$

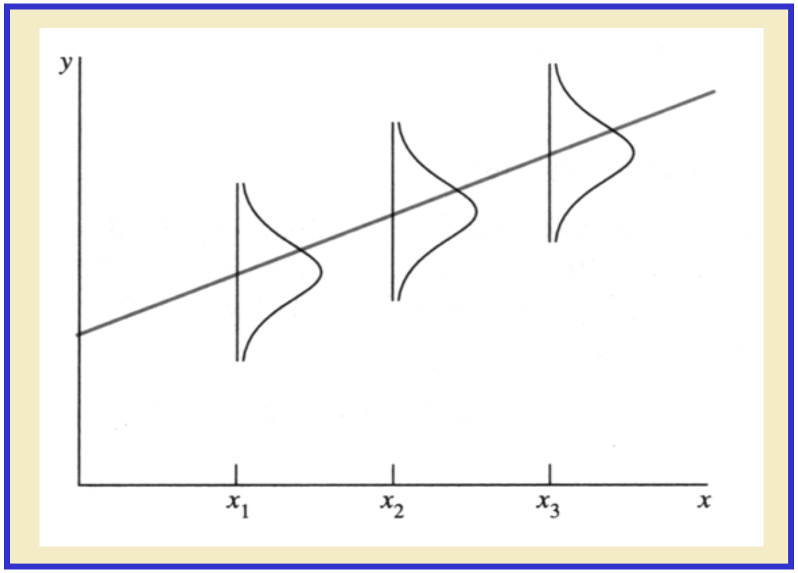


Points deviate from the line of means by an amount ε where ε has a normal distribution with mean 0 and variance σ^2.

The Random Error

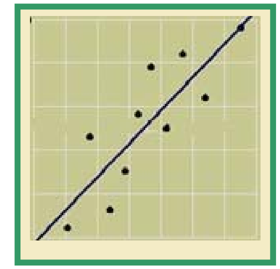


- The line of means, $E(y) = \mathbf{a} + \mathbf{b}x$, describes average value of y for any fixed value of x.
- The population of measurements is generated as y deviates from the population line by **e**. We estimate **a** and **b** using sample information.



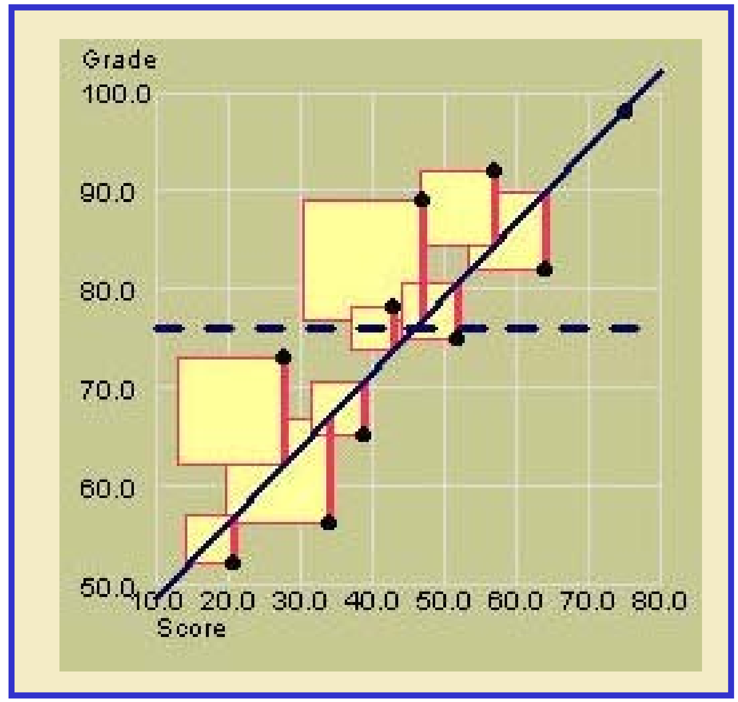
MY APPLET

The Method of Least Squares



- The equation of the best-fitting line is calculated using a set of n pairs (x_i, y_i).

- We choose our estimates a and b to estimate a and b so that the vertical distances of the points from the line, are minimized.

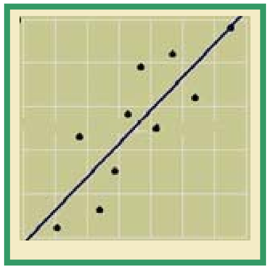


$$\text{Bestfitting line}: \hat{y} = a + bx$$

$$\text{Choose } a \text{ and } b \text{ to minimize}$$

$$\text{SSE} = \sum (y - \hat{y})^2 = \sum (y - a - bx)^2$$

Least Squares Estimators



Calculate the sums of squares:

$$S_{xx} = \sum x^2 - \frac{(\sum x)^2}{n} \qquad S_{yy} = \sum y^2 - \frac{(\sum y)^2}{n}$$

$$S_{xy} = \sum xy - \frac{(\sum x)(\sum y)}{n}$$

Best fitting line: $\hat{y} = a + bx$ where

$$b = \frac{S_{xy}}{S_{xx}} \quad \text{and} \quad a = \bar{y} - b\bar{x}$$

Example

The table shows the math achievement test scores for a random sample of n = 10 college freshmen, along with their final calculus grades.

Student	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Math test, x	39	43	21	64	57	47	28	75	34	52
Calculus grade, y	65	78	52	82	92	89	73	98	56	75

Use your calculator to find the sums and sums of squares.

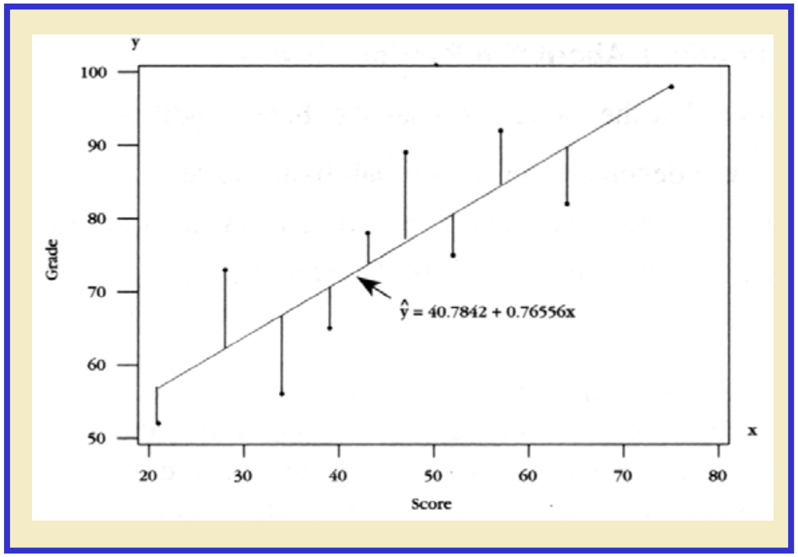
$$\sum x = 460 \qquad \sum y = 760$$

$$\sum x^2 = 23634 \quad \sum y^2 = 59816$$

$$\sum xy = 36854$$

$$\bar{x} = 46 \qquad \bar{y} = 76$$

Example



$$S_{xx} = 23634 - \frac{(460)^2}{10} = 2474$$

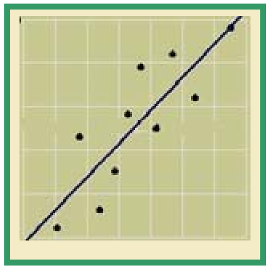
$$S_{yy} = 59816 - \frac{(760)^2}{10} = 2056$$

$$S_{xy} = 36854 - \frac{(460)(760)}{10} = 1894$$

$$b = \frac{1894}{2474} = .76556 \quad \text{and} \quad a = 76 - .76556(46) = 40.78$$

$$\text{Best fitting line:} \quad \hat{y} = 40.78 + .77x$$

The Analysis of Variance



- The total variation in the experiment is measured by the **total sum of squares**:

$$\text{Total SS} = S_{yy} = \sum (y - \bar{y})^2$$

The Total SS is divided into two parts:

- SSR (sum of squares for regression): measures the variation explained by using *x* in the model.
- SSE (sum of squares for error): measures the leftover variation not explained by *x*.

The Analysis of Variance

We calculate

$$SSR = \frac{(S_{xy})^2}{S_{xx}} = \frac{1894^2}{2474}$$

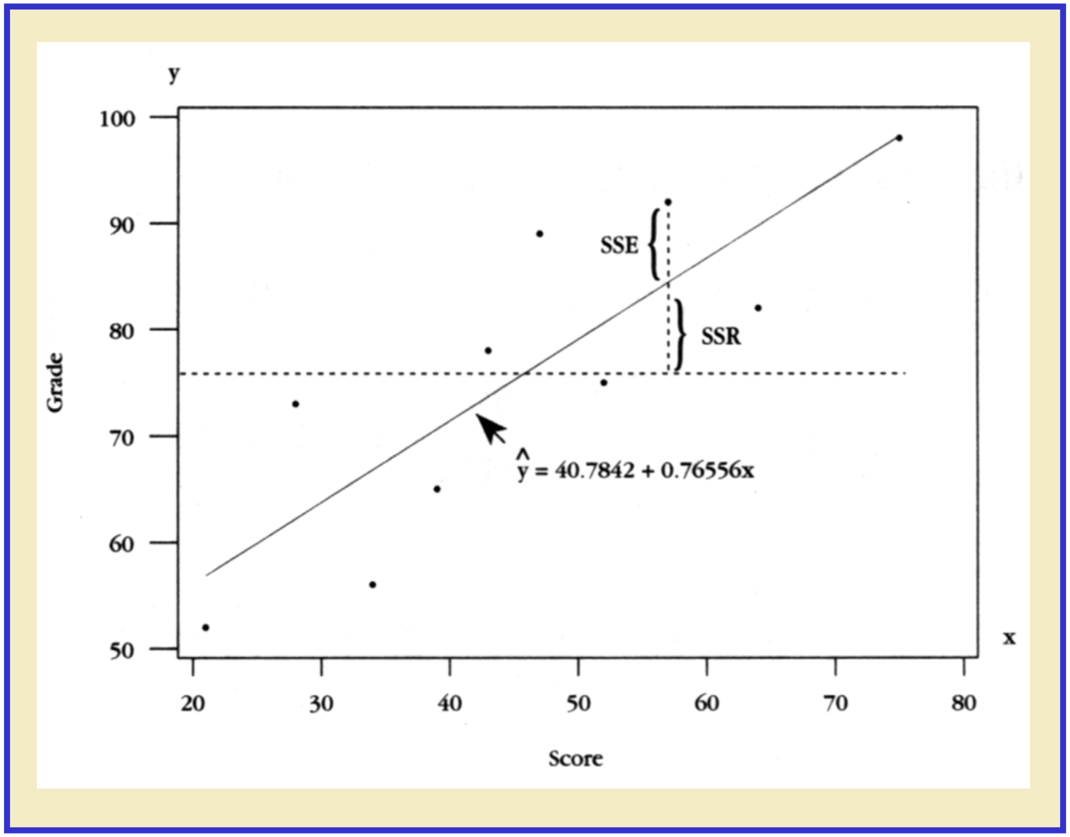
$$= 1449.9741$$

$$SSE = \text{Total SS} - SSR$$

$$= S_{yy} - \frac{(S_{xy})^2}{S_{xx}}$$

$$= 2056 - 1449.9741$$

$$= 606.0259$$



The ANOVA Table

Total df = $n - 1$

Regression df = 1

Error df = $n - 1 - 1 = n - 2$

Mean Squares

MSR = SSR/(1)

MSE = SSE/(n-2)

Source	df	SS	MS	F
Regression	1	SSR	SSR/(1)	MSR/MSE
Error	$n - 2$	SSE	SSE/(n-2)	
Total	$n - 1$	Total SS		

The Calculus Problem

$$SSR = \frac{(S_{xy})^2}{S_{xx}} = \frac{1894^2}{2474} = 1449.9741$$

$$SSE = \text{Total SS} - SSR = S_{yy} - \frac{(S_{xy})^2}{S_{xx}}$$

$$= 2056 - 1449.9741 = 606.0259$$

Source	df	SS	MS	F
Regression	1	1449.9741	1449.9741	19.14
Error	8	606.0259	75.7532	
Total	9	2056.0000		

Minitab Output

To test $H_0 : \beta = 0$

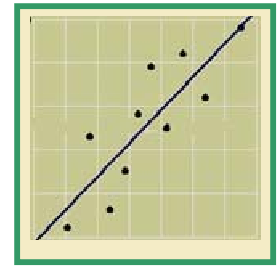
Least squares regression line

```
Regression Analysis: y versus x
The regression equation is y = 40.8 + 0.766 x
Predictor          Coef     SE Coef      T        P
Constant          40.784     8.507     4.79     0.001
x                 0.7656     0.1750    4.38     0.002

S = 8.70363    R-Sq = 70.5%    R-Sq(adj) = 66.8%

Analysis of Variance
Source           DF     SS       MS       F        P
Regression        1   1450.0   1450.0   19.14    0.002
Residual Error    8    606.0     75.8
Total             9   2056.0
```

$\sqrt{MSE}$

Regression coefficients, *a* and *b*

$t^2 = F$



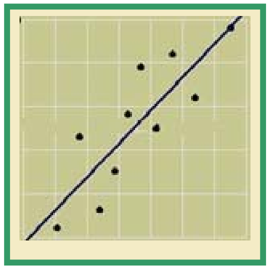
Measuring the Strength of the Relationship

- If the independent variable *x* is of useful in predicting *y*, you will want to know how well the model fits.
- The strength of the relationship between *x* and *y* can be measured using:

$$\text{Correlation coefficient}: r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$$

$$\text{Coefficient of determination}: r^2 = \frac{S_{xy}^{\ 2}}{S_{xx}S_{yy}} = \frac{\text{SSR}}{\text{Total SS}}$$

Measuring the Strength of the Relationship



- Since Total SS = SSR + SSE, r^2 measures
- the proportion of the total variation in the responses that can be explained by using the independent variable x in the model.
- the percent reduction the total variation by using the regression equation rather than just using the sample mean y-bar to estimate y.

For the calculus problem, $r^2 = .705$ or 70.5%. The model is working well!

$$r^2 = \frac{\text{SSR}}{\text{Total SS}}$$

Logistic Regression

Logistic Regression

$$Y=f(x)$$

- Logistic or “Logit” regression investigates the relationship between response variables (Y’s) and one or more predictor variables (X’s) where:
  - **Y’s are categorical, not continuous**
  - X’s can be either continuous or categorical

Logistic Regression

- Both logistic regression and least squares regression investigate the relationship between a response variable and one or more predictors.

- A practical difference between them is that logistic regression techniques are used with categorical response variables, and linear regression techniques are used with continuous response variables.

MINITAB provides three logistic regression procedures that you can use to assess the relationship between one or more predictor variables and a categorical response variable of the following types:

Variable type	Number of categories	Characteristics	Examples
Binary	2	two levels	success, failure yes, no
Ordinal	3 or more	natural ordering of the levels	none, mild, severe fine, medium, coarse
Nominal	3 or more	no natural ordering of the levels	blue, black, red, yellow sunny, rainy, cloudy

Example

In some situations, Six Sigma practitioners find a *Y* that is discrete and *X*s that are continuous. How can a regression equation be developed in these cases? Black Belt training indicated that the correct technique is something called logistic regression. An example about a well-known space shuttle accident can help to demystify logistic regression using the simplest logistic regression – binary logistic regression, where the *Y* has just two potential outcomes (i.e., “yes” or “no,” or 0 or 1).

- The data in Table 1 comes from the Presidential Commission on the Space Shuttle Challenger Accident (1986). The data consists of the number of the flight, the air temperature at the time of the launch and whether or not there was damage to the booster rocket field joints (no = 0, yes = 1).

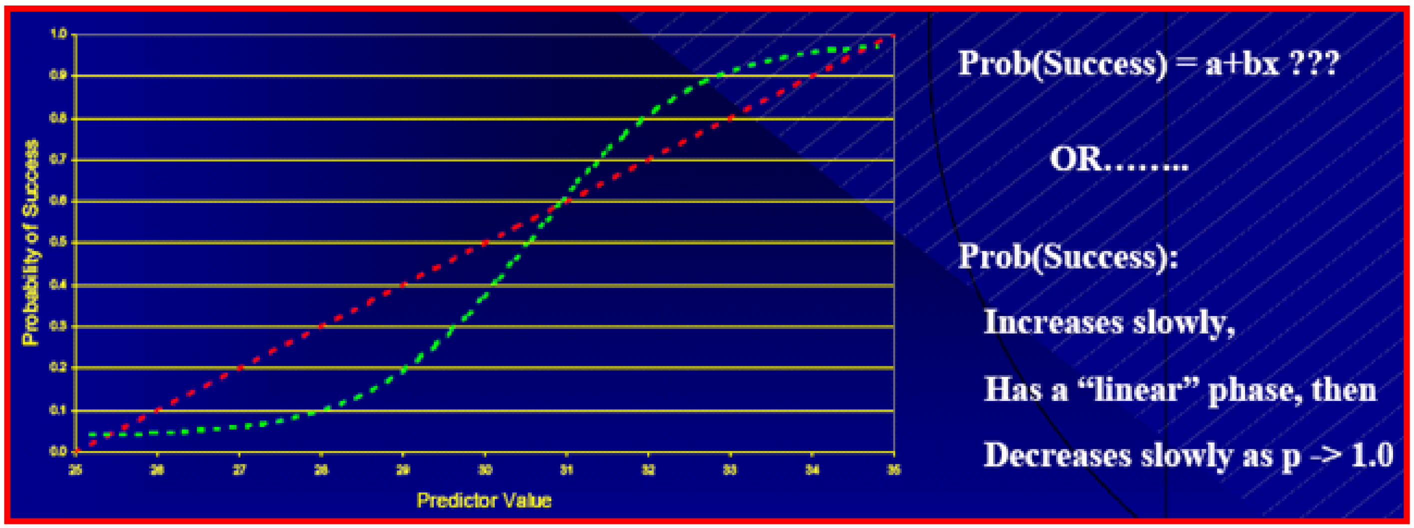
Flight	Temp.	Damage	Flight	Temp.	Damage
STS 1	66	0	STS 51A	67	0
STS 2	70	1	STS 51C	53	1
STS 3	69	0	STS 51D	67	0
STS 5	68	0	STS 51B	75	0
STS 6	67	0	STS 51G	70	0
STS 7	72	0	STS 51F	81	0
STS 8	73	0	STS 51I	76	0
STS 9	70	0	STS 51J	79	0
STS 41B	57	1	STS 61A	75	1
STS 41C	63	1	STS 61B	76	0
STS 41D	70	1	STS 61C	58	1

Formulate the Regression Model

- Any regression requires a continuous output or Y. However, in this case the Y is discrete with only two categories or two events: Damage – yes or no. What to do? The "trick" behind the logistic regression is to turn the discrete output into a continuous output by calculating the probability (p) for the occurrence of a specific event. That means, the logistic regression provides a model to predict the p for a specific event for Y (here, the damage of booster rocket field joints, $p = P[Y=1]$) given any value of X (here, the temperature at the time of the launch). The logistic regression equation has the form:
- This function is the so-called "logit" function where this regression has its name from. The procedure for modeling a logistic model is determining the actual percentages for an event as a function of the X and finding the best constant and coefficients fitting the different percentages.
- This is exactly the equation that comes out of statistical software's output for logistics regression:

Why Logistic Regression?

- Binomial data violates normality, equal variance assumption
  - $\mu_k = n * p$ $\sigma_k^2 = n * p * (1-p) = \mu_k * (1-p)$
  - Variance changes as the mean changes
  - The relationship between p, the likelihood of "success" and the predictor variables might not be linear



Binary Logistic Regression

- We will demystify logistic regression using the simplest logistic regression – binary logistic regression (where the Y has just two potential outcomes, i.e., "yes" or "no," or 0 or 1)

- These events are often described as success or failure

- For each possible values for the independent (X) variables, there is a probability that a “success” occurs

The linear logistic model fitted by maximum likelihood is:

- $Y = b_0 + b_1{*}X_1 + b_2{*}X_2 \ldots\ldots + b_k{*}X_k$
- Where Y = logit transformation of the odds based on p = Prob(event)
  - Odds = $\left(\dfrac{p}{1-p}\right)$ Logit = $\ln\left(\dfrac{p}{1-p}\right)$

Deriving Probability from Logit Results

$$\ln\left(\frac{p}{1-p}\right) = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \ldots..$$

$$\left(\frac{p}{1-p}\right) = e^{b_0+b_1x_1+b_2x_2+b_3x_3+\ldots..} = \text{"}odds\text{"}$$

$$p = (1-p) \times e^{b_0+b_1x_1+b_2x_2+b_3x_3+\ldots..}$$

$$p = e^{b_0+b_1x_1+b_2x_2+b_3x_3+\ldots..} - pe^{b_0+b_1x_1+b_2x_2+b_3x_3+\ldots..},$$

$$p(1+e^{b_0+b_1x_1+b_2x_2+b_3x_3+\ldots..}) = e^{b_0+b_1x_1+b_2x_2+b_3x_3+\ldots..}$$

$$p = \frac{e^{b_0+b_1x_1+b_2x_2+b_3x_3+\ldots..}}{1+e^{b_0+b_1x_1+b_2x_2+b_3x_3+\ldots..}} = \frac{\text{"}odds\text{"}}{1+\text{"}odds\text{"}}$$

Key Concepts

- So far, we have not used any statistical tool to prioritize X's.

- Depending upon the data characteristics of Y & X, we can choose the appropriate tool

Y \ X	Continuous	Discrete
Continuous	Correlation & Regression	Hypothesis testing
Discrete	Logistic Regression	Chi-square

Lean Six Sigma Black Belt

by
Rajiv Purkayastha
Six sigma MBB

Improve Phase Overview

What is the Improve phase?

The Improve phase is when your team:

- Selects those product performance characteristics that must be improved to achieve the improvement goal by identifying the major sources of variation in the process.
- Develops and pilots process improvements

Generate and Select Solutions

Possible Paths For Solution Development

Out-of-the-Box Thinking OR Structured DOE Approach

Generate New
Solution Ideas

Investigate
Specific Xs

Both approaches require experimentation
Both approaches result in one or more good solution options

Generate and Select Solutions

Tools used for Generating Solution

- Brain Storming
- Creative Thinking
- 6-3-5 Brainwriting

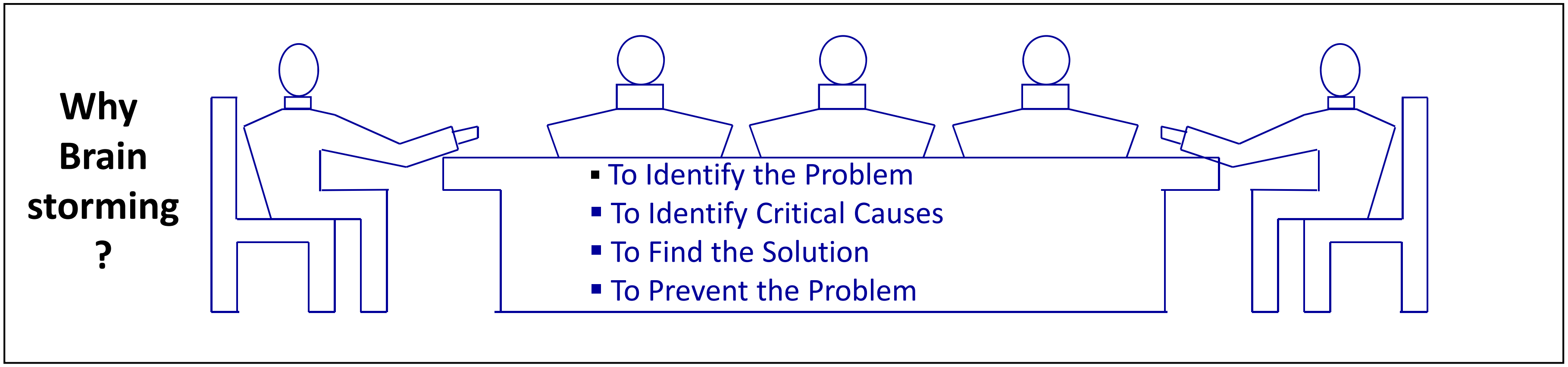
Brainstorming

Brain Storming

Definition -

Brain storming is a Team approach to generate creative ideas in a short time

Brain storming plays an important role to build a **Cause and Effect Diagram**



Types of Brain Storming -

- **Free Wheeling : Spontaneous flow of ideas by all team members**
- **Round Robin : Team members take turns suggesting ideas**
- **Card Method : Team members write ideas on cards with no discussion**

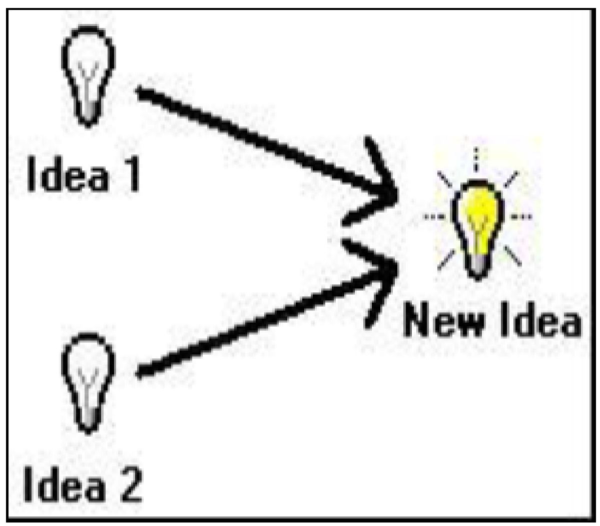
Brain Storming

Brain storming can be conducted in the way-

- Every person in a group must give an idea as their turn arises.
- Forces even shy people to participate.
- Creates a certain amount of pressure to contribute.

Creative Thinking

- Creative thinking techniques work to stimulate original ideas.
- New ideas happen when two or more ideas are accidentally or deliberately merged.
- Creative thinking techniques provide the method for **deliberately** combining ideas.



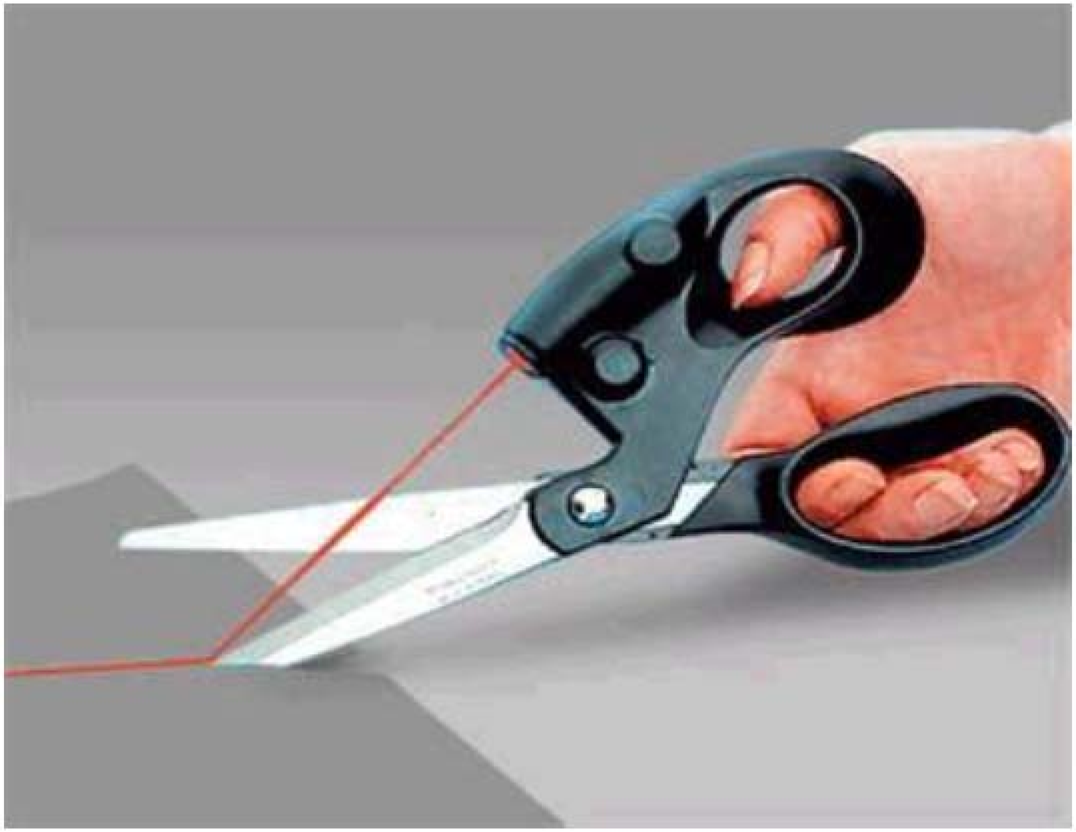
Innovative creations that make life easier

Cup &
Cookies

Innovative creations that make life easier

Banana
Guard

Innovative creations that make life easier



Laser
Scissors

Innovative creations that make life easier

"The Thing"-
Infant Pillow

Ergonomic infant pillow designed by a mom to mimic the size, weight, touch and feel of her hand and forearm to help her baby with comfort

Innovative creations that make life easier

Wheel-
Moving Bench

Innovative creations that make life easier

Day
Clock

6-3-5 Brainwriting

6-3-5 Brainwriting or Method 635

- **6-3-5 Brainwriting** (also known as the 6-3-5 Method, or Method 635) is a group creativity technique originally developed by Professor Bernd Rohrbach in 1968.

- Based on the concept of Brainstorming, the aim of 6-3-5 Brainwriting is to generate 108 new ideas in half an hour.

- In a similar way to brainstorming, it is not the quality of ideas that matters but the quantity.

Brainwriting vs. Brainstorming

- **The difference with Brainstorming is that in Brainwriting each participant thinks and records ideas individually, without any verbal interaction.**

- Several studies (notably Diehl and Strobe's, from 1987 to 1994) tested brainstorming teams extensively and realized that **participants working in isolation consistently outperformed participants working in groups, both in quantity and quality of ideas generated.**

Steps for Brainwriting

- The technique involves 6 participants who sit in a group and are supervised by a moderator. Each gets a sheet of paper with the same problem statement written at the top.

- Each participant thinks up 3 (unedited) ideas every 5 minutes. The difference from brainstorming is that the ideas are being recorded in private.

- At the end of 5 minutes each participant passes the sheet of paper to the participant to the left.

- Each participant now reads the ideas that were previously written and a new five-minute round starts. Each participant must again come up with three new ideas. Participants are free to use the ideas already on the sheet as triggers — or to ignore them altogether.

Steps for Brainwriting

- After 6 rounds in 30 minutes the group has thought up a total of 108 ideas.
- After the idea-gathering phase is completed, the ideas are read, discussed and consolidated with the help of the moderator

Prioritize Solutions

Selecting Top Solutions

- Criteria-Based Matrix or Grid analysis
- Design Of Experiments (DOE)

Criteria-Based Matrix

Where to use CBM

- More than a few criteria for solution selection
- Solutions scoring differently under each criteria
- The weight for each criteria differs

How to use CBM

- Identify Criteria and assign weightage
- The Team members to give votes to each idea
- The votes are then multiplied with the weights of the criteria established for selection.
- The total scores obtained at the bottom of the solution are used to select the same

Criteria-Based Matrix Template

Criteria Based Matrix						
			Likely Solutions			
S. No	Criteria	Weight	Solution A		Solution B	
			Score	Wtd Score	Score	Wtd Score
1	List each Criterion here					
2	XXXXXXXXXXXXXXXXXXXXX					
3	XXXXXXXXXXXXXXXXXXXXX					
4	XXXXXXXXXXXXXXXXXXXXX					
5	XXXXXXXXXXXXXXXXXXXXX					
6	XXXXXXXXXXXXXXXXXXXXX					
7	XXXXXXXXXXXXXXXXXXXXX					
8	XXXXXXXXXXXXXXXXXXXXX					
9	XXXXXXXXXXXXXXXXXXXXX					
10	XXXXXXXXXXXXXXXXXXXXX					
	Total					

CBT Template

Process For Criteria-Based Matrix Process

- Record a final list of solutions
- Screen against musts
- Create a list of "want" criteria
- Weight the list of "want" criteria
- Compare the list of solutions to the weighted criteria
- Tally and discuss total scores for each solution

- *All solutions that have been screened for acceptability can now be examined via the criteria- based matrix.*
- *Unlike the "musts" criteria, "want" criteria are used to compare the relative benefits of different solutions. "Want" criteria can include items such as: less training, lower cost, brief implementation, etc.*
- *Once the list of "want" criteria is generated, the top "want" is identified and labeled as a 10. The rest of the "wants" are ranked relative to that "want."*
- *The solutions are then compared to each "want." The solution that best fulfills that "want" is ranked "10" and the other solutions are ranked relative to that solution.*

Example…Criteria-Based Matrix

Imagine that you are buying a house. You have already decided what your "musts" criteria are and they include budget (how much can you spend) and size (square footage).

Although you have looked at many houses, only two meet the "must" criteria.

Now you are ready to compare these two houses on the want criteria you have already established.

Criteria Based Matrix						
			Likely Solutions			
S. No	Criteria	Weight	Solution A		Solution B	
			Score	Wtd Score	Score	Wtd Score
1	Big Yard	2				
2	Neighborhood With Kids	8				
3	Good Schools	10				
4	Proximity To Work	9				
5	Three-Car Garage	6				
6						
7						
8						
9						
10						
	Total			0		0

Example…Criteria-Based Matrix

After examining the list, you decide the "top want" is good schools and you ranked that "want" a 10. The rest of the criteria are ranked relative to that "want".

The next step is to compare the two houses against each specific criteria. The house that BEST meets the criteria is ranked "10" and the other house is ranked relative to that 10.

After both houses have been ranked on all criteria, multiply the weight by the score for each criteria to get the weighted score for each criteria. Total the scores and you will be able to then discuss the results of the matrix.

Criteria Based Matrix

			Likely Solutions			
S. No	Criteria	Weight	Solution A		Solution B	
			Score	Wtd Score	Score	Wtd Score
1	Big Yard	2	4	8	10	20
2	Neighborhood With Kids	8	10	80	5	40
3	Good Schools	10	10	100	8	80
4	Proximity To Work	9	9	81	10	90
5	Three-Car Garage	6	10	60	4	24
6						
7						
8						
9						
10						
	Total			329		254

Design of Experiments (DOE)

- Definition
  - A test in which purposeful changes are made to the certain parameters of a system so that one may observe and quantify the changes in the outputs.
- Origin
  - 1920's with Sir R Fisher
  - Has an agriculture based nomenclature, e.g. treatments

Purpose of DOE

- To Study or compare the effect of a factor
- To determine the important factor
- Optimization- to determine the setting of some factors to minimize an output
- To create a mathematical relationship between factors and outputs.

Learning Process

EVENT

- BREAK DOWNS,
- CUSTOMER COMPLAINTS
- ACCIDENT
- INCREASE IN COSTS
- DECREASE IN SALES

RECOGNIZE

- QUALITY INDICATOR
- SERVICE DEFECT RATE
- FLC
- SPC

RESEARCH

IMPROVEMENT PROCESS

- To be able to improve processes, we need to find the causes affecting outputs and analyze their behavior. There are two ways to do this:
  - **<u>Continuous Control and Observation</u>**

  This is the familiar method. In most of the cases, we let the problem appear itself and when it is recognized, we try to find the causes that has changed the natural behavior. Statistical process control is used for this.
  - **<u>Design of Experiments</u>**

  In this method, to gain more information the levels of the probable inputs are manipulated and their responses are checked.

Design of experiments involve activities carried out to obtain informative results.

Experiment Methods

Problem: Decreasing the fuel consumption of a car or increasing the kms by unit liter

Output:10 km / lt → to 15 km/lt

- **Trial and error**
- **One factor at a time**
- **Design of Experiments**
  - **Full Factorial Experiments**
  - **Fractional Factorial Experiments**
  - **Response Surface Methods**

Trial and Error

Problem: Fuel consumption of a car

10 km/lt → 15 km/lt

Factors:

- Change the brand of the fuel.
- Drive slowly
- Increase tire pressure
- Change spark plugs
- Increase tire diameter
- Clean the car

What if it works?

What if it does not work?

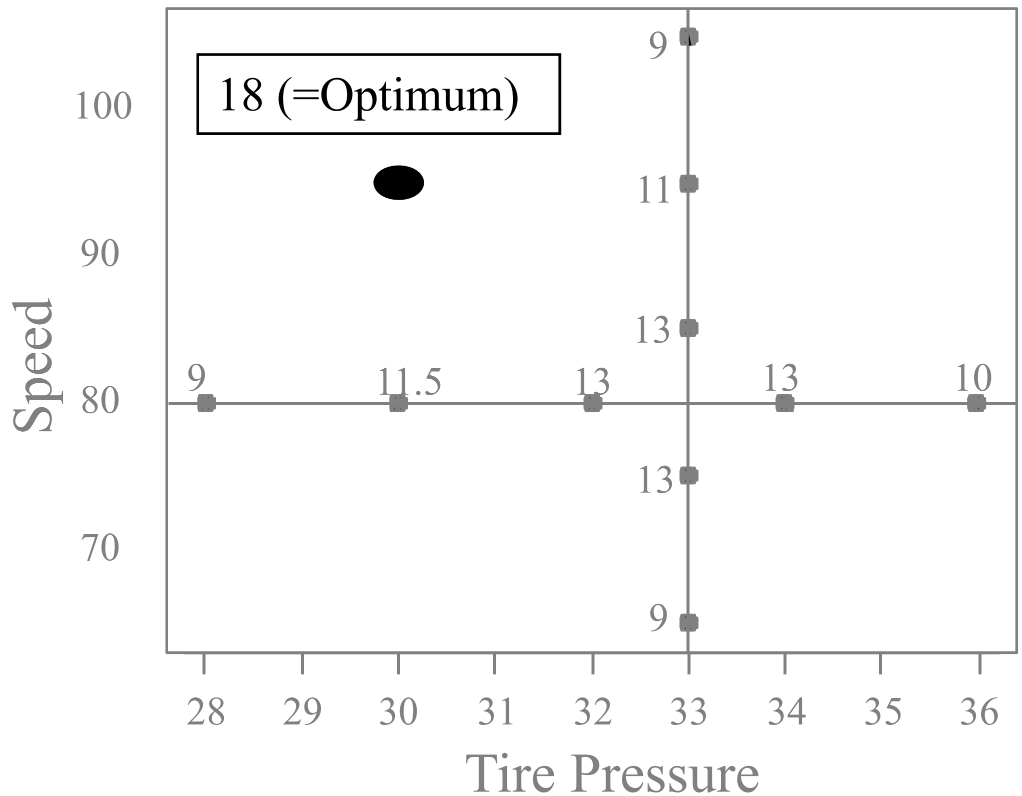
One Factor at a Time

Problem: Decreasing the fuel consumption of a car from 10 km/lt to 15 km/lt.

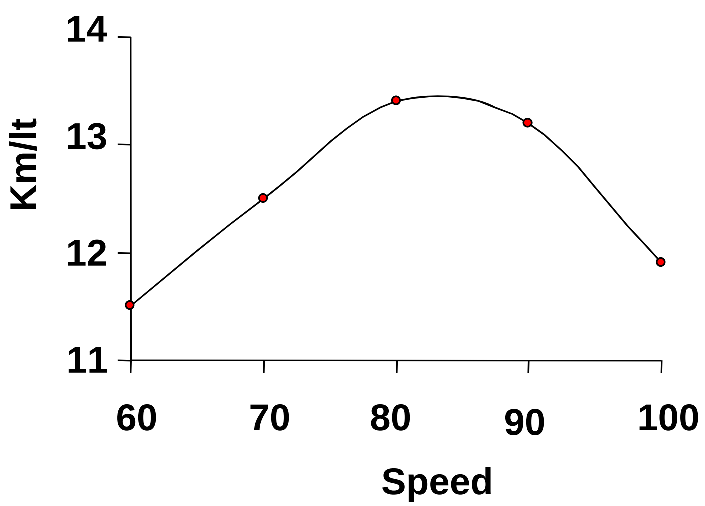
Speed (A)	Octane Ratio (B)	Tire. Press. (C)	km/lit (Y)
80	85	30	12,5
100	85	30	11,5
80	91	30	13,5
80	91	35	13,5

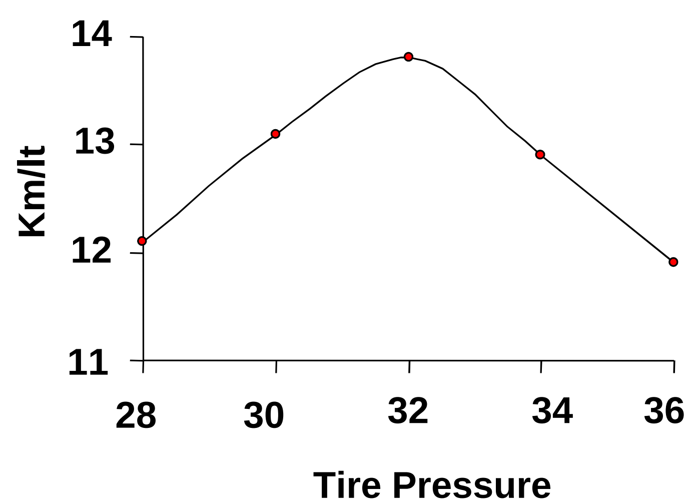
How many trials ?...
How much information ?...

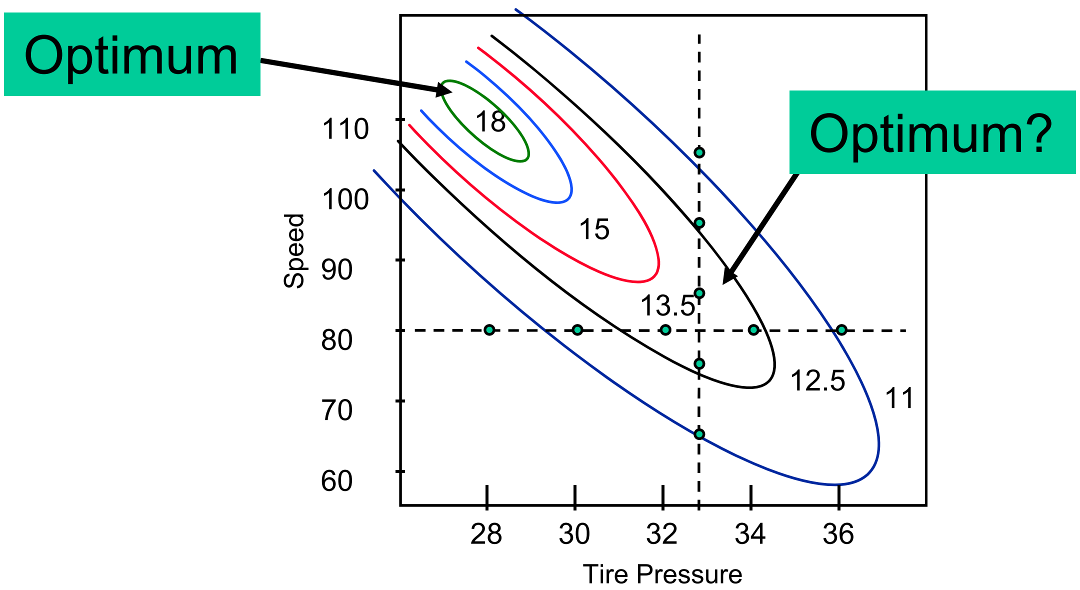
One Factor at a Time



One Factor at a Time

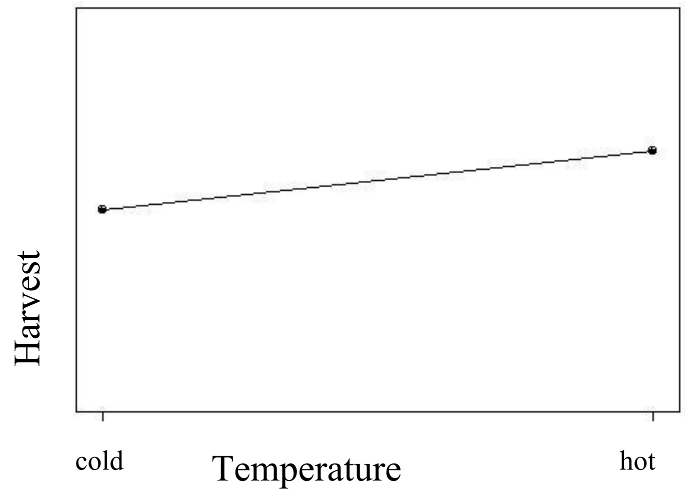




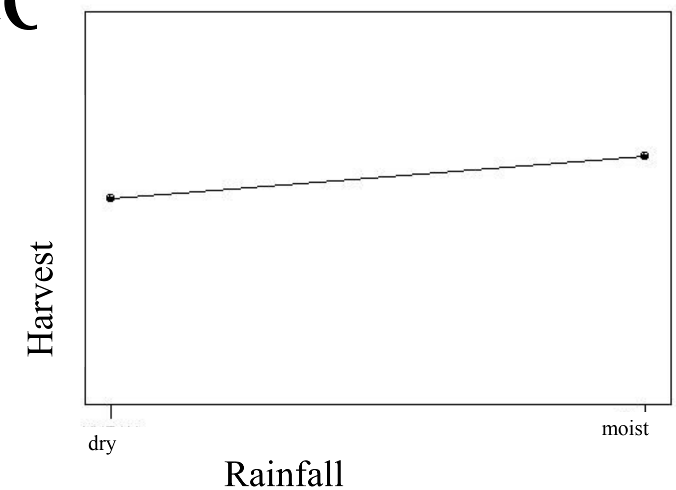


Optimumum cannot be found with this method.

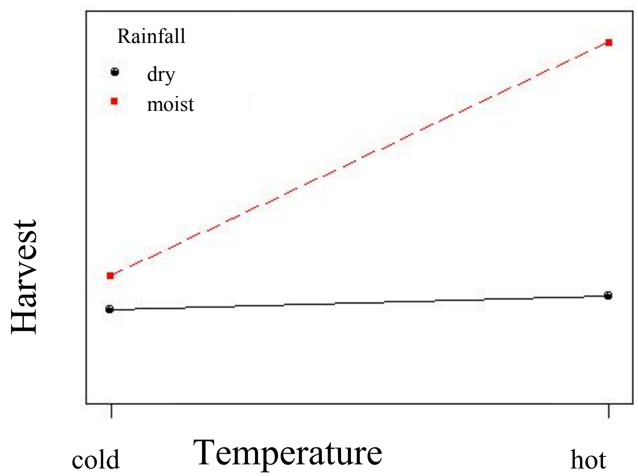
Based on temperature



Based on rainfall



Annual Comparison



**** One factor at a time approach does not give information about interactions***

Full Factorial Experiments

Speed (A)	Octane (B)	Tire Pressure (C)	km/lit (Y)
80	85	30	Y_1
100	85	30	Y_2
80	91	30	Y_3
100	91	30	Y_4
80	85	35	Y_5
100	85	35	Y_6
80	91	35	Y_7
100	91	35	Y_8

How many trials?
How many observation at each trial?

Full Factorial Experiments

ADVANTAGES

- Gives information about all interactions
- More effective than one factor at a time experiments
- Could be easily planned and analyzed
- Can be done for both 2 or more than two-level factors
- Valid both for quantitative and qualitative factors

LIMITATIONS

- If there are a lot of factors and levels, combinations will increase and eventually experiment will take much more time
- Could be waste of time if wrong factors or levels are selected
- Could only be used with quantitative outputs

DOE Terminology

RESPONSE – Outcome that are obtained from experimental units after treatments have been applied, also called as dependent variable.

FACTOR- A factor is one of the controlled or uncontrolled variables whose influence upon the response is being studied in the experiment. Factors are also known as the X's.

e.g. temperature, pressure etc.

LEVEL - The "levels" of a factor are the values of the factor being examined in the experiment. For quantitative factors, each chosen value becomes a level, e.g., if the experiment is to be conducted at two different temperatures, then the factor temperature has two "levels".

REPEATION- This is running the experiment twice on each trial combination, without changing the setting, i.e. no other run in between

DOE Terminology

REPLICATION- This is running the experiment twice on each trial combination, but with a change of setting, i.e. some other run in between. It is done to reduce the impact of inherent variation in the process.

RANDOMIZATION- Runs are made in random order as opposed to a standard order to avoid lurking variables that change over time. This is to eliminate the effect of lurking variable, uncontrolled factor.

Full Factorial DOE

- Used when there are many factors of interest

- If there are ‘a’ factors and ‘b’ levels, ‘ba’ combinations are possible.

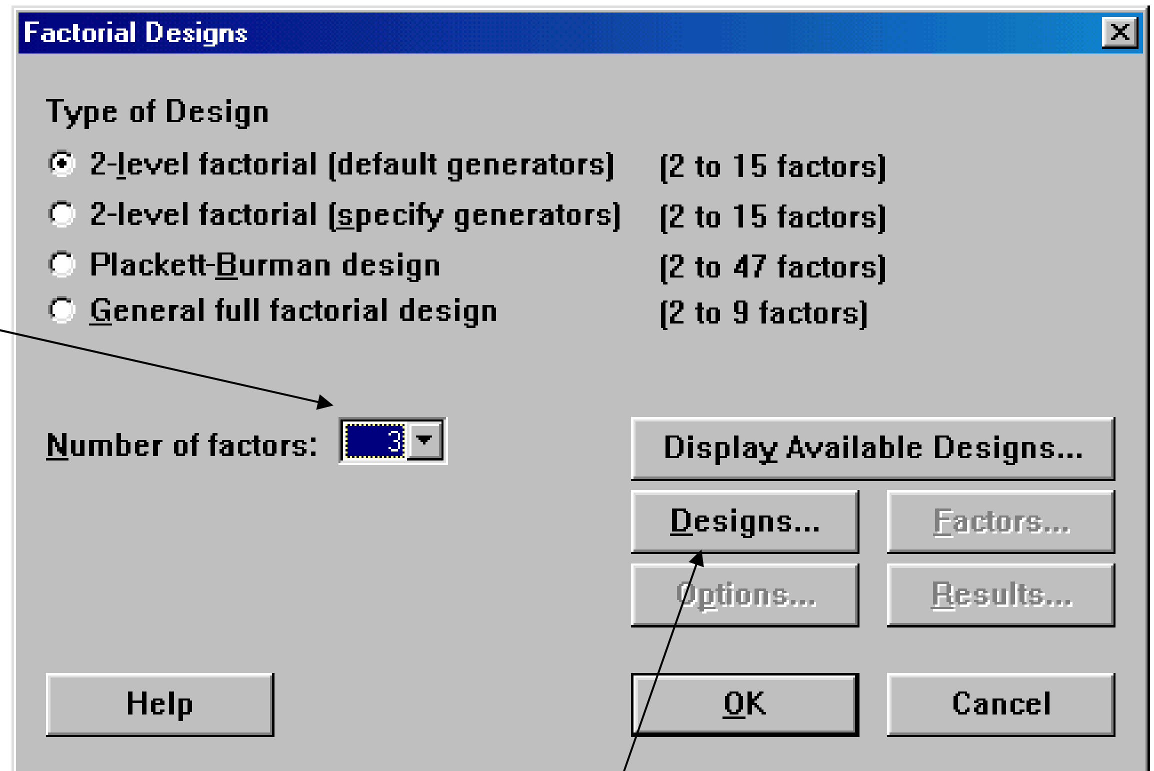
Example : 3 Factors & 2 levels

Factors	**Level**	
	-1 (low)	**+ (High)**
Temperature	**100 ^{0}C**	**120 ^{0} C**
Pressure	**2 Kg/Cm^2**	**5 Kg/Cm^2**
Catalyst	**A**	**B**

Creating Designs in Minitab

- STAT > DOE > CREATE FACTORIAL DESIGN

3 factors relating to
Temperature,
Time &
Concentration



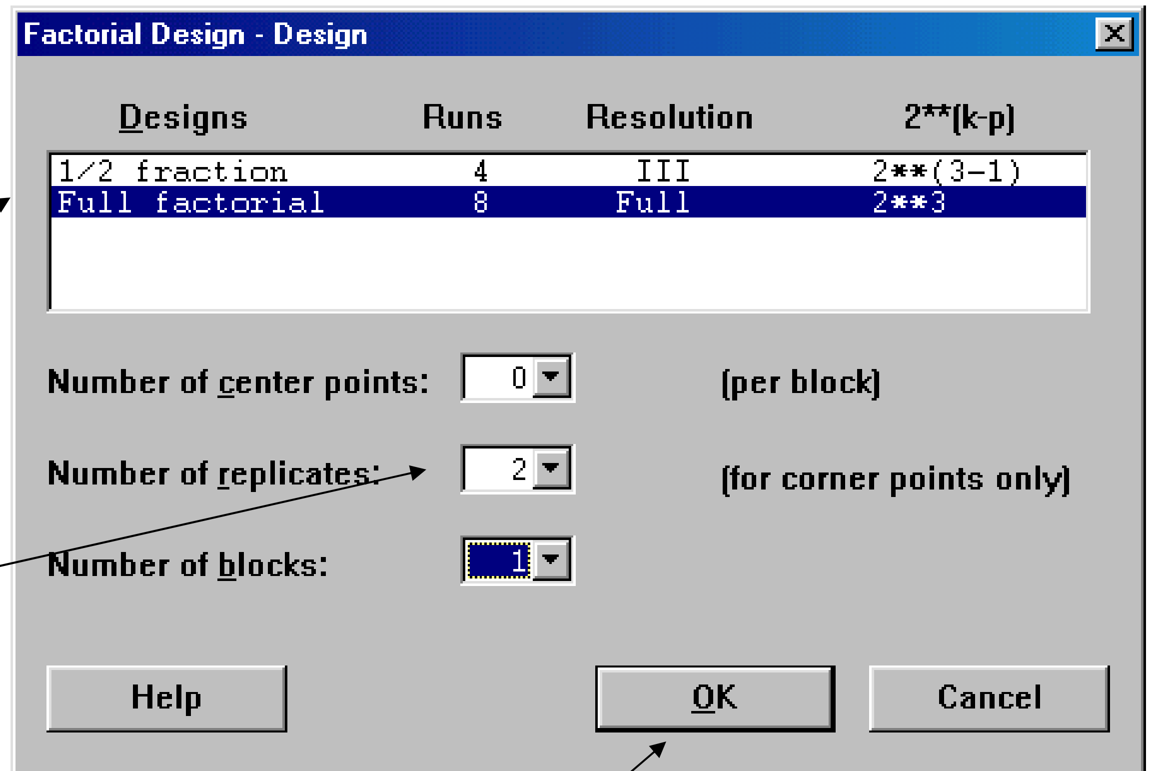
Click on Designs

Creating Designs in Minitab

➢ STAT > DOE > CREATE FACTORIAL DESIGN > DESIGNS

Choose a full factorial
or fractional factorial design

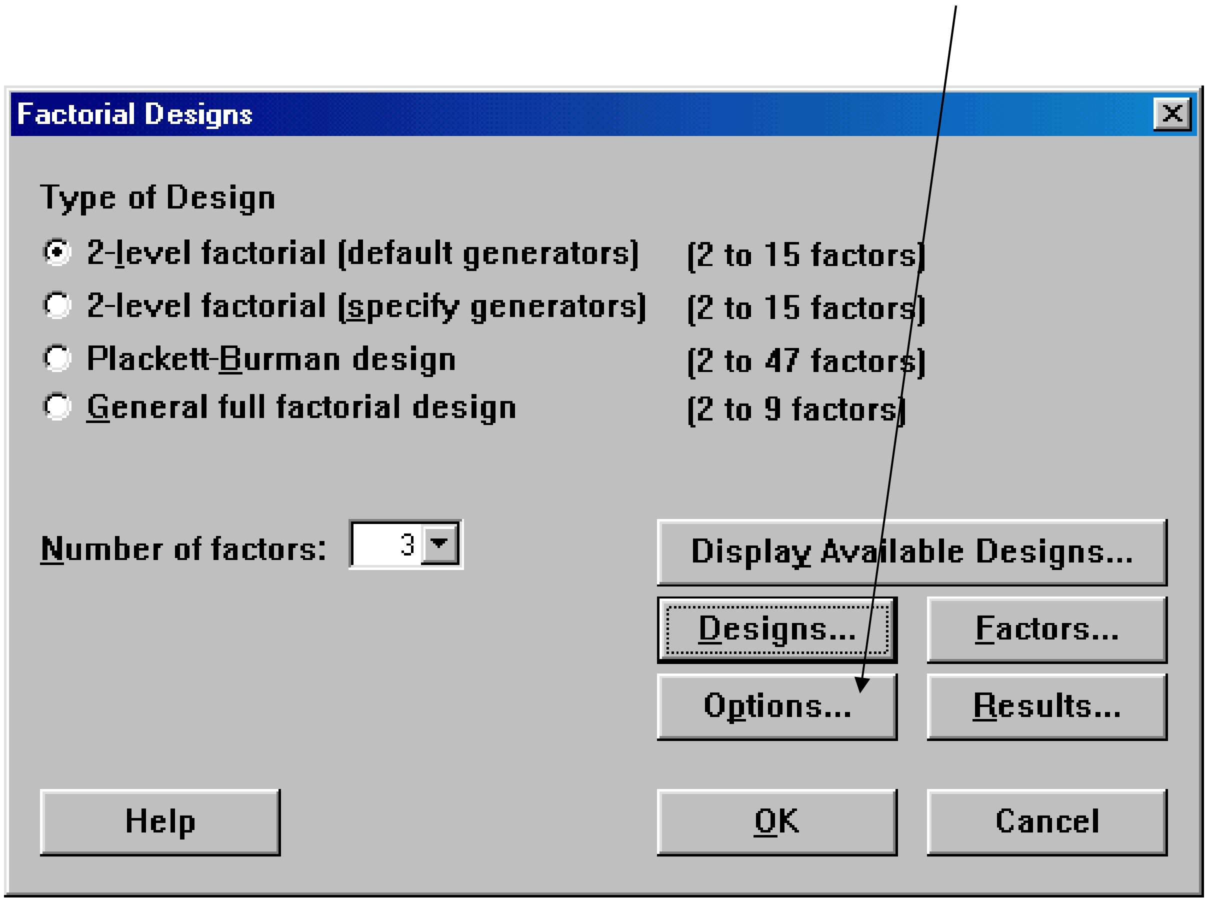
Choose two replicates



Click on OK

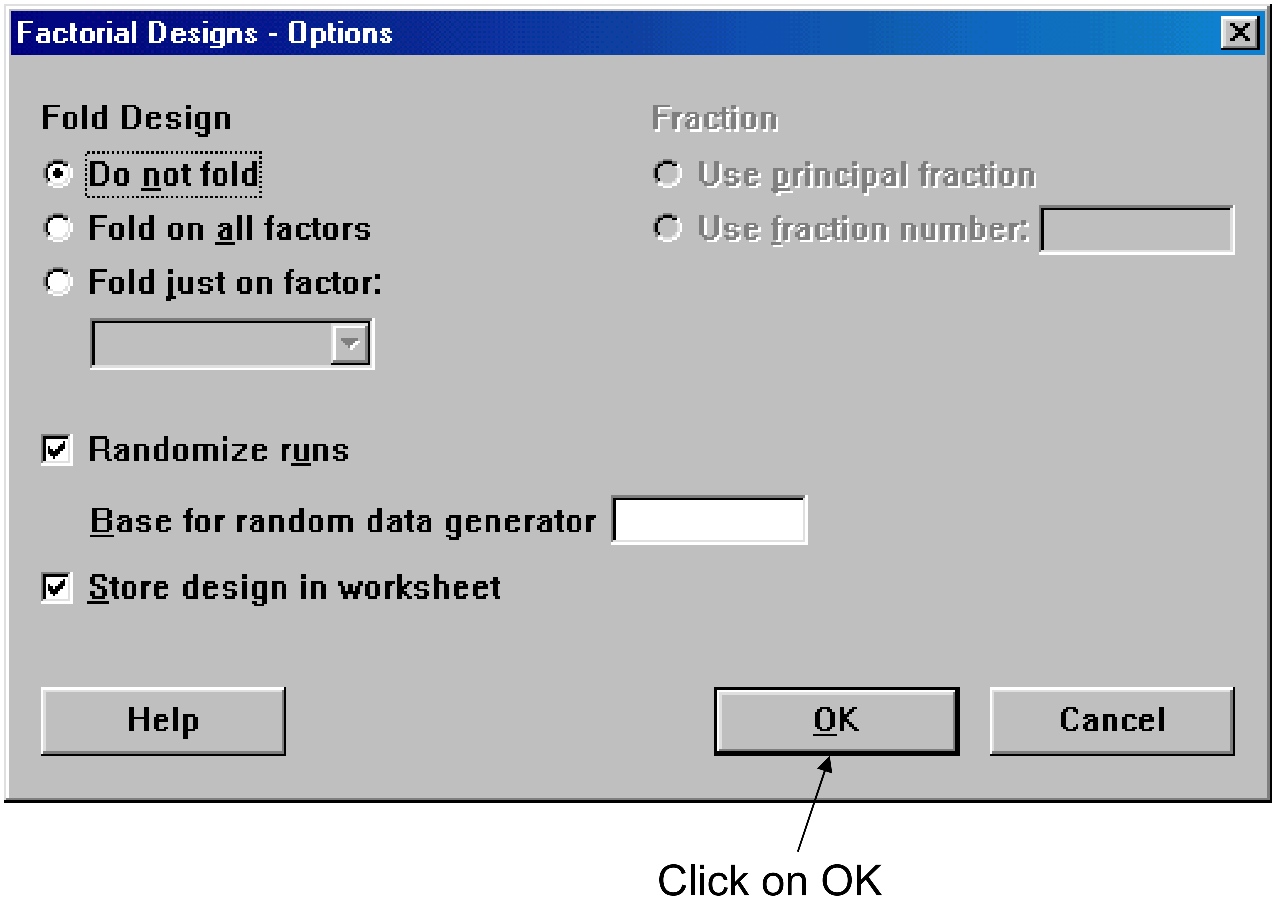
Creating Designs in Minitab

- STAT > DOE > CREATE FACTORIAL DESIGN > Click on options



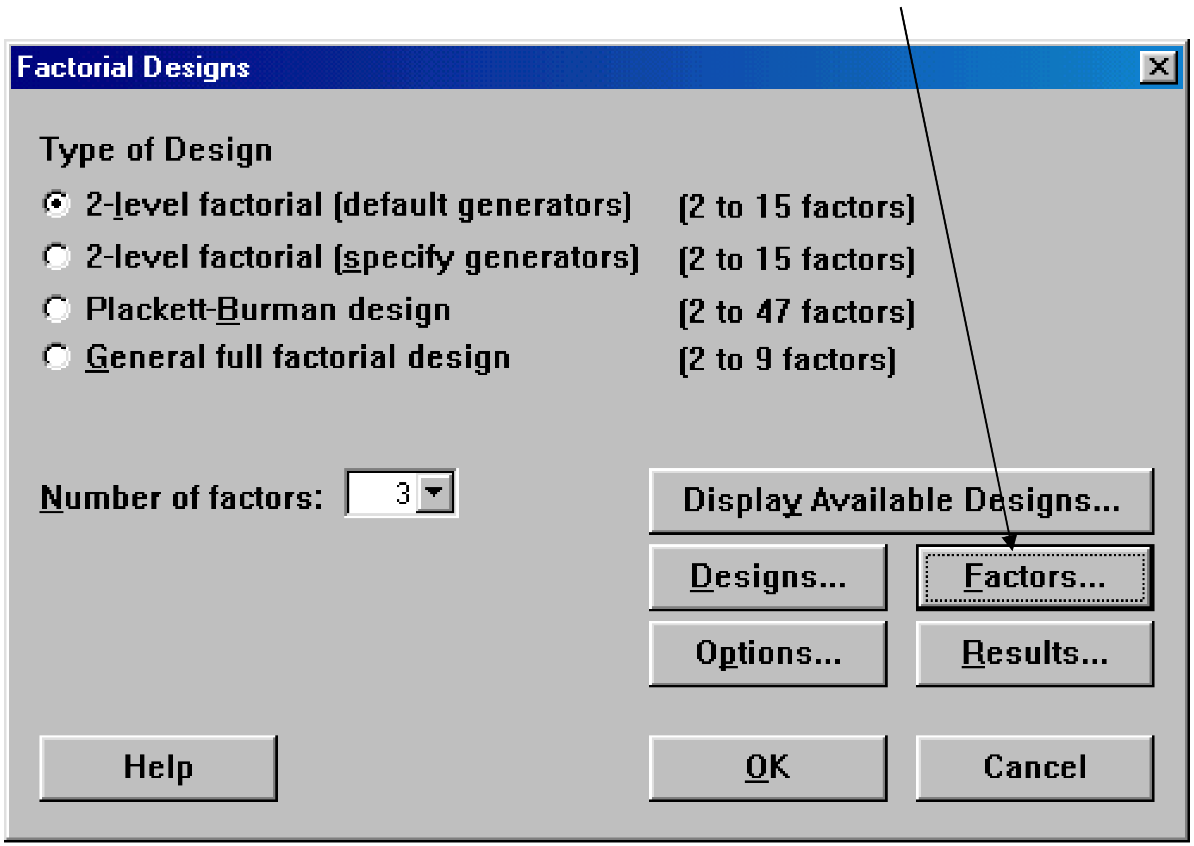
Creating Designs in Minitab

➢ STAT > DOE > CREATE FACTORIAL DESIGN > OPTIONS



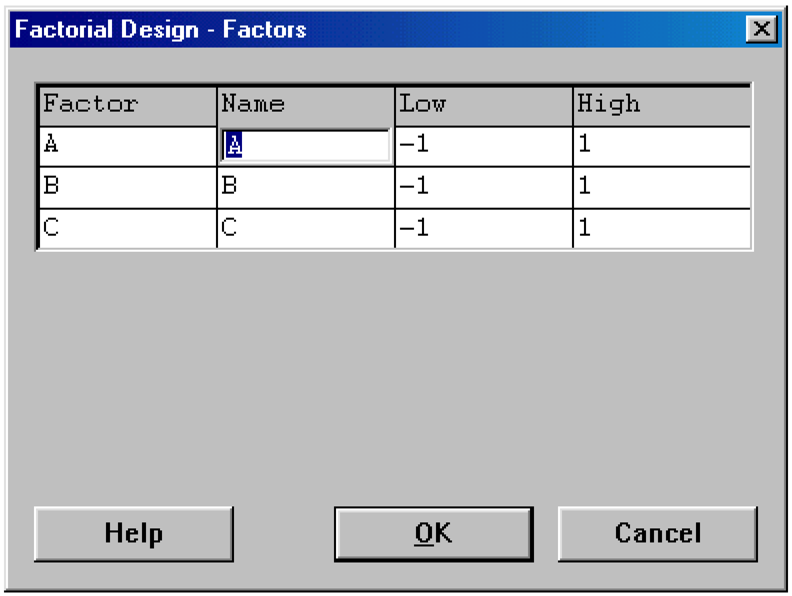
Creating Designs in Minitab

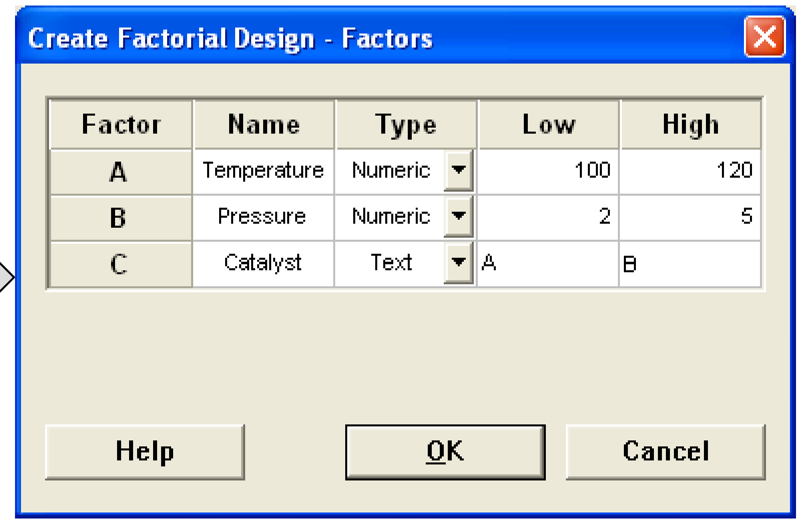
- STAT > DOE > CREATE FACTORIAL DESIGN > Click on factors



Creating Designs in Minitab

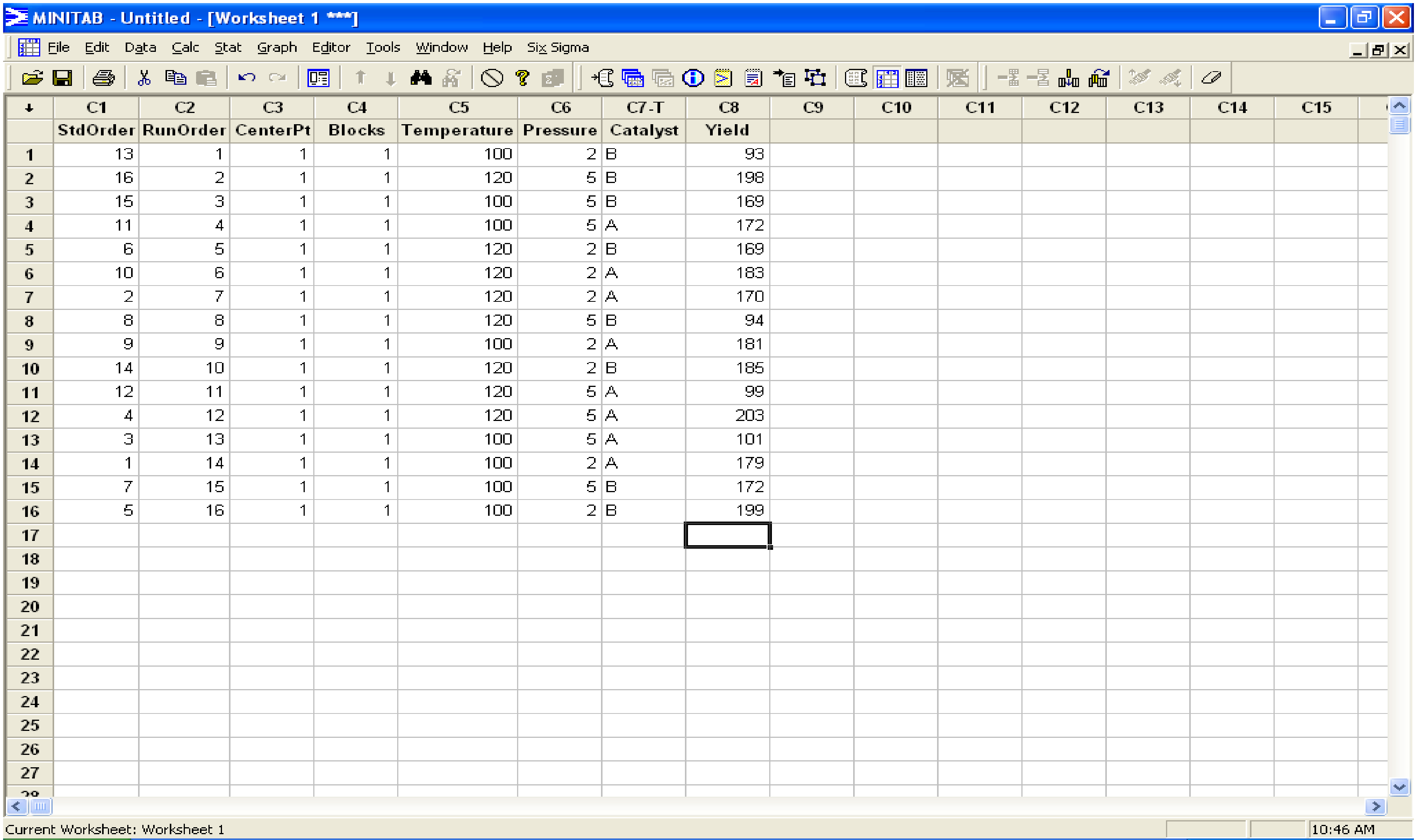
➢ STAT > DOE > CREATE FACTORIAL DESIGN > FACTORS





Click n OK

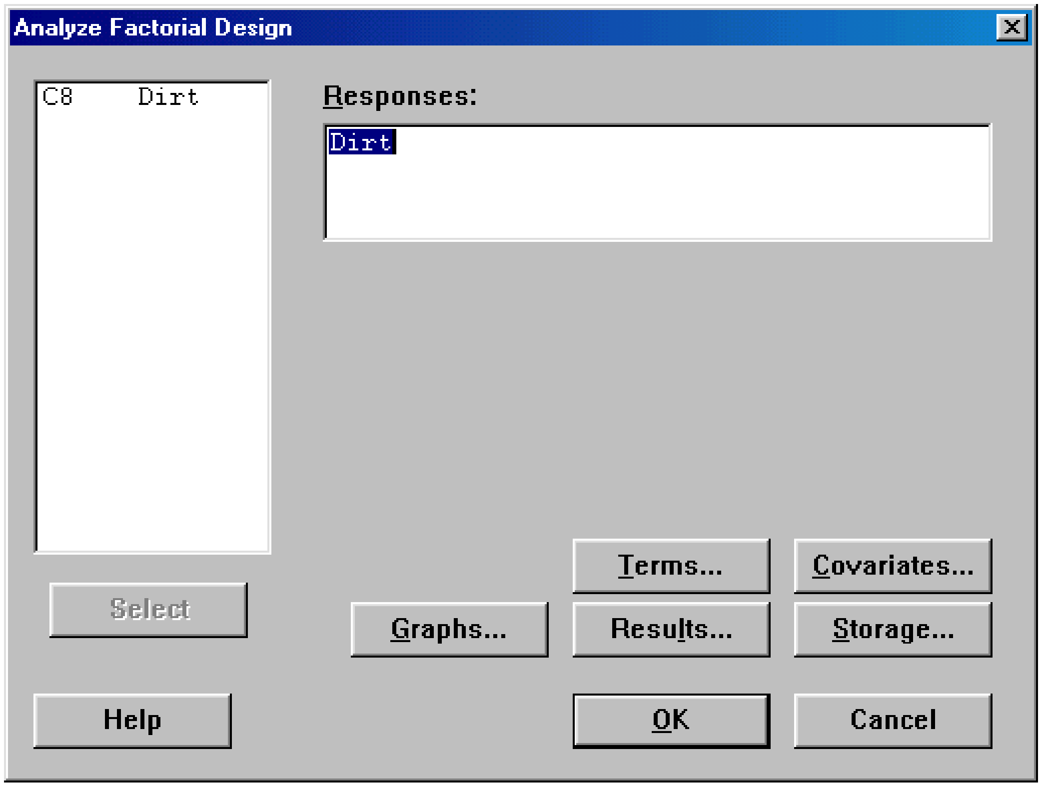
Creating Designs in Minitab



↓	C1	C2	C3	C4	C5	C6	C7-T	C8
	StdOrder	RunOrder	CenterPt	Blocks	Temperature	Pressure	Catalyst	Yield
1	13	1	1	1	100	2	B	93
2	16	2	1	1	120	5	B	198
3	15	3	1	1	100	5	B	169
4	11	4	1	1	100	5	A	172
5	6	5	1	1	120	2	B	169
6	10	6	1	1	120	2	A	183
7	2	7	1	1	120	2	A	170
8	8	8	1	1	120	5	B	94
9	9	9	1	1	100	2	A	181
10	14	10	1	1	120	2	B	185
11	12	11	1	1	120	5	A	99
12	4	12	1	1	120	5	A	203
13	3	13	1	1	100	5	A	101
14	1	14	1	1	100	2	A	179
15	7	15	1	1	100	5	B	172
16	5	16	1	1	100	2	B	199

Creating Designs in Minitab

- STAT > DOE > ANALYZE FACTORIAL DESIGN



- ‘Analyze factorial design’ will be enabled only if Minitab was used to create the design

Creating Designs in Minitab

- Following is the Minitab Output:

Fractional Factorial Fit

Estimated Effects and Coefficients for **Dirt (coded units)**

Term	Effect	Coef	StDev Coef	T	P
Constant		49.500	0.6903	71.70	0.000
Temp	-12.000	-6.000	0.6903	-8.69	0.000
Time	-6.750	-3.375	0.6903	-4.89	0.001
Conc	0.250	0.125	0.6903	0.18	0.861
Temp*Time	6.750	3.375	0.6903	4.89	0.001
Temp*Conc	0.750	0.375	0.6903	0.54	0.602
Time*Conc	2.500	1.250	0.6903	1.81	0.108
Temp*Time*Conc	-2.500	-1.250	0.6903	-1.81	0.108

Creating Designs in Minitab

- Following is the Minitab Output:

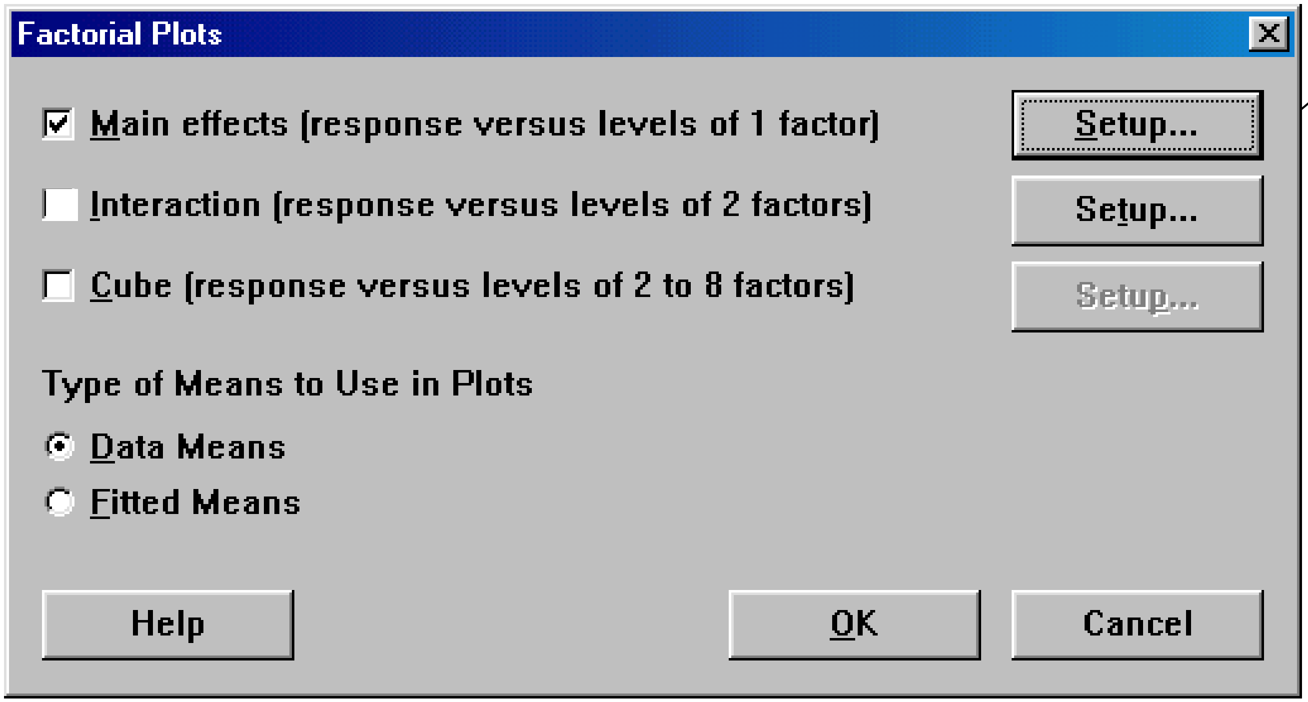
Analysis of Variance for Dirt (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	758.50	758.500	252.833	33.16	0.000
2-Way Interactions	3	209.50	209.500	69.833	9.16	0.006
3-Way Interactions	1	25.00	25.000	25.000	3.28	0.108
Residual Error	8	61.00	61.000	7.625		
Pure Error	8	61.00	61.000	7.625		
Total	15	1054.00				

Interpreting Results

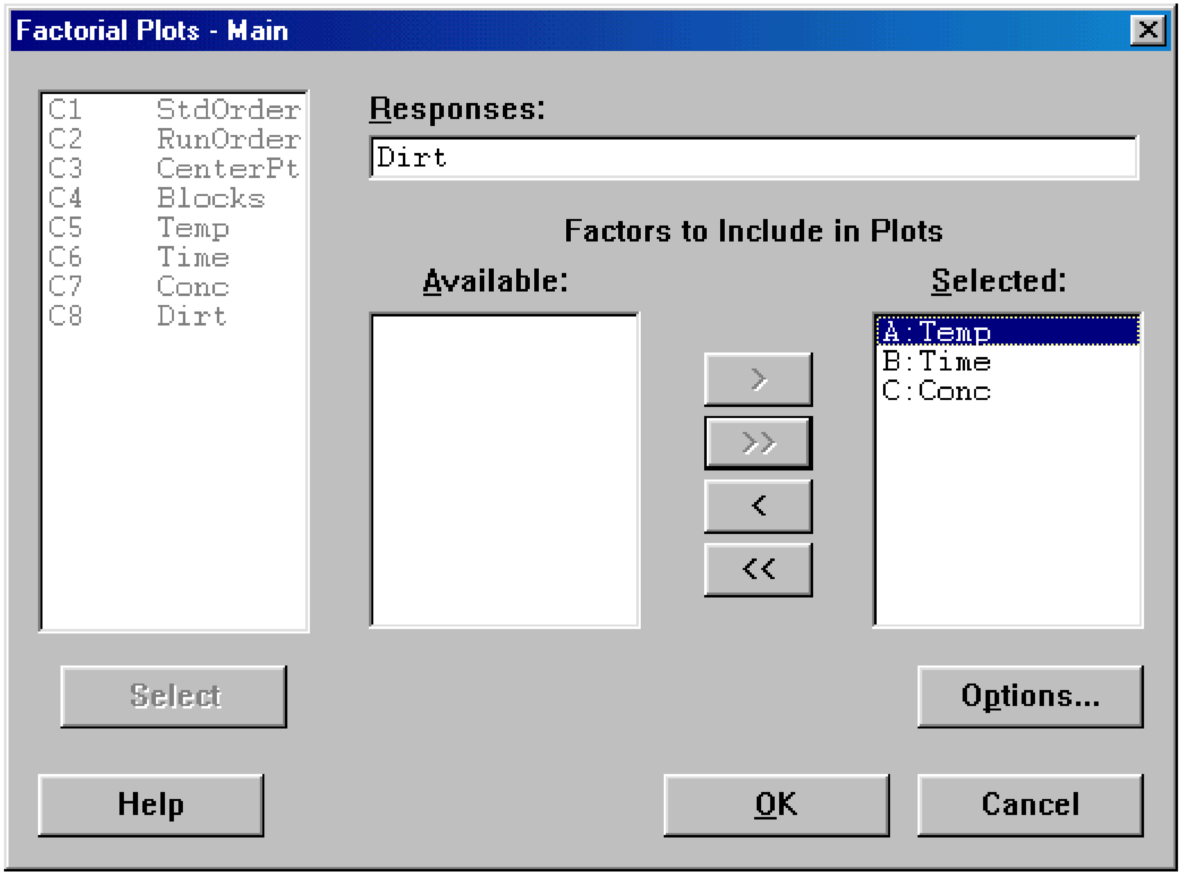
- STAT > DOE > FACTORIAL PLOTS

Click on Setup



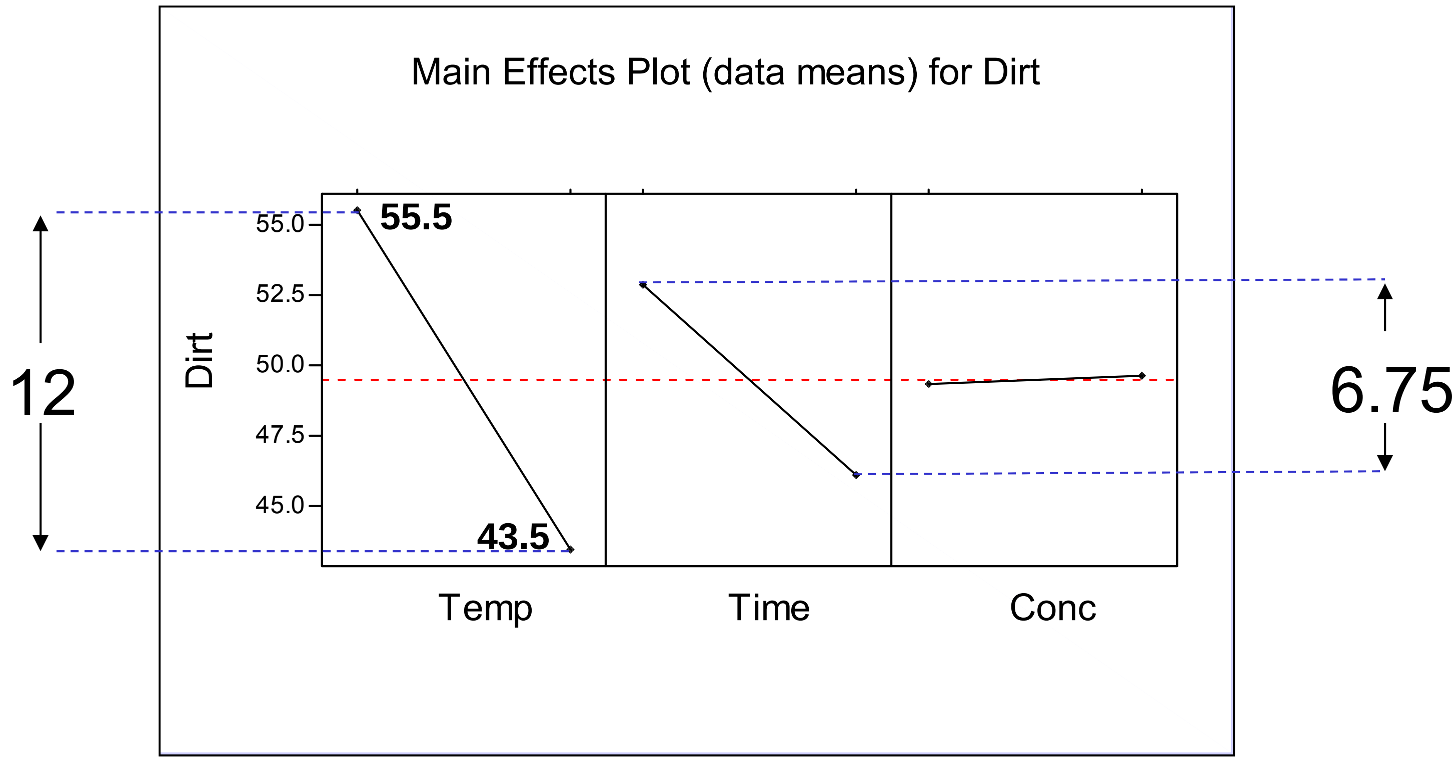
Interpreting Results

- STAT > DOE > FACTORIAL PLOTS > Main effects plot



Interpreting Results

- STAT > DOE > FACTORIAL PLOTS > Main effects plot

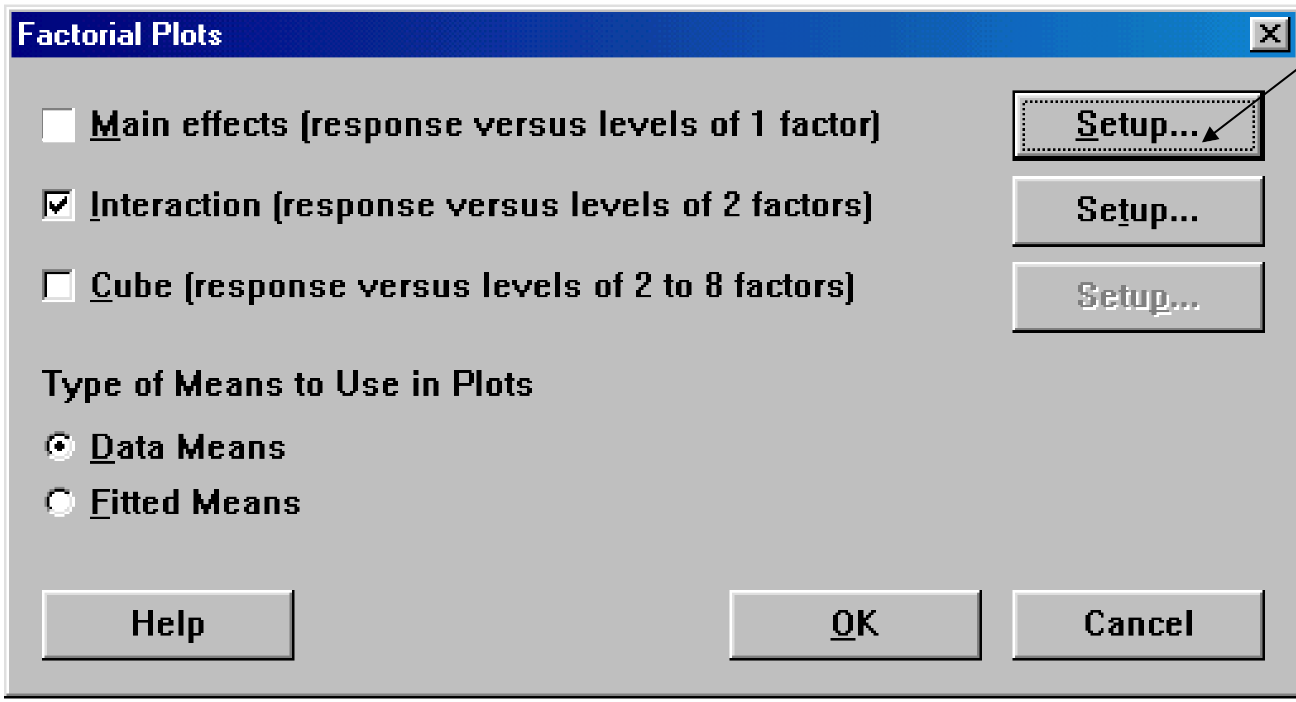


- It's clear that temperature has the greatest effect, time has a moderate effect & concentration has the least effect

Interpreting Results

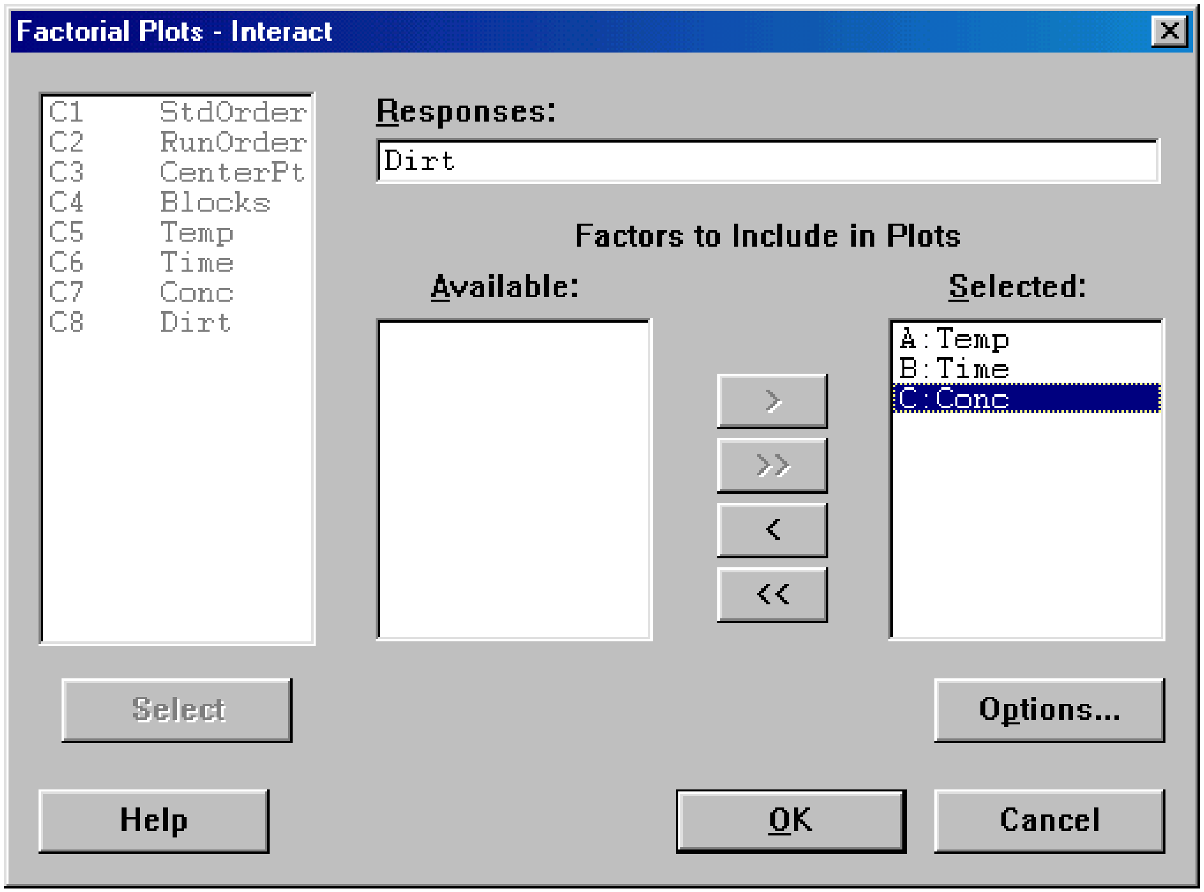
- STAT > DOE > FACTORIAL PLOTS

Click on Setup



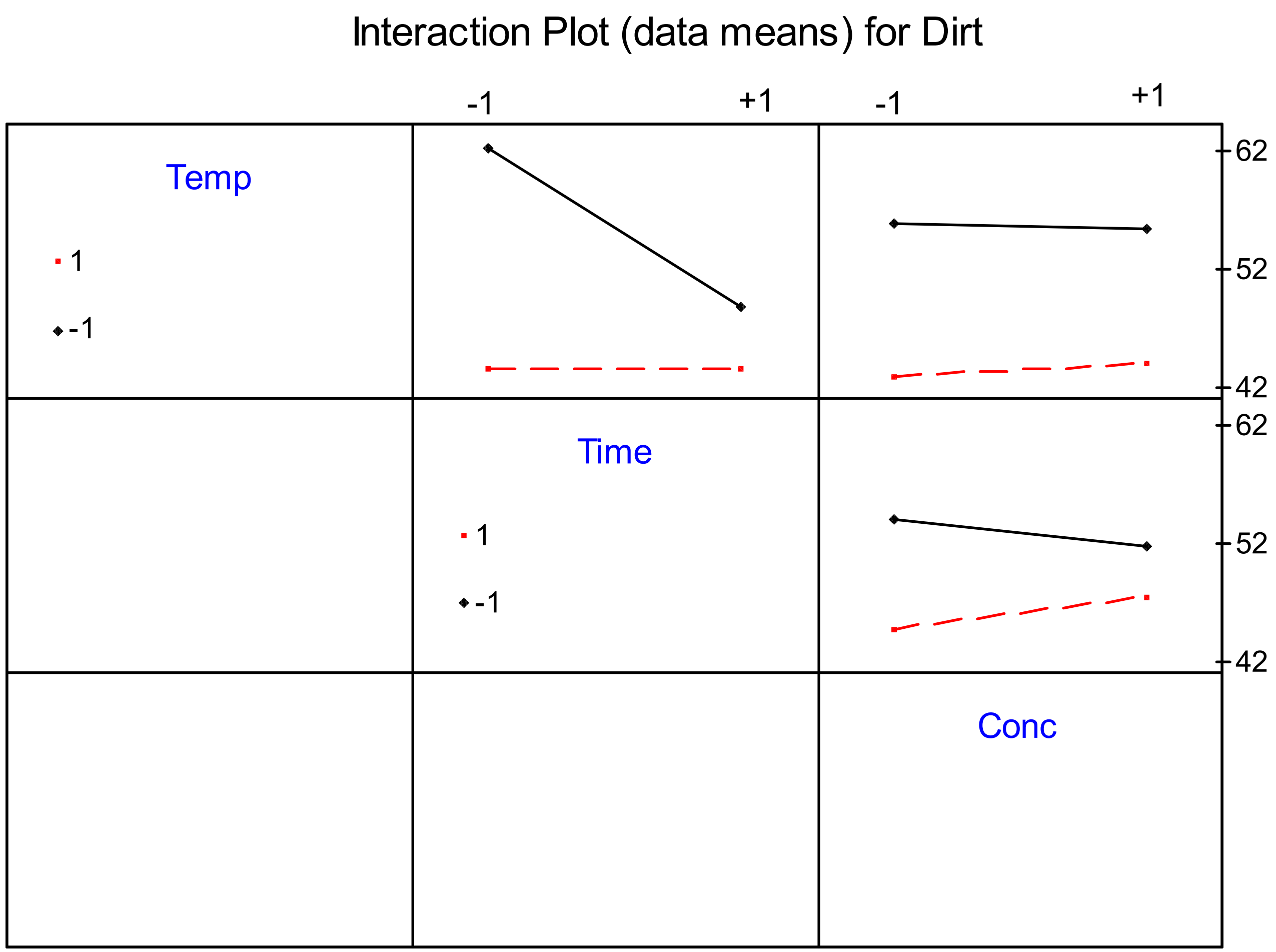
Interpreting Results

- STAT > DOE > FACTORIAL PLOTS > Interaction plot



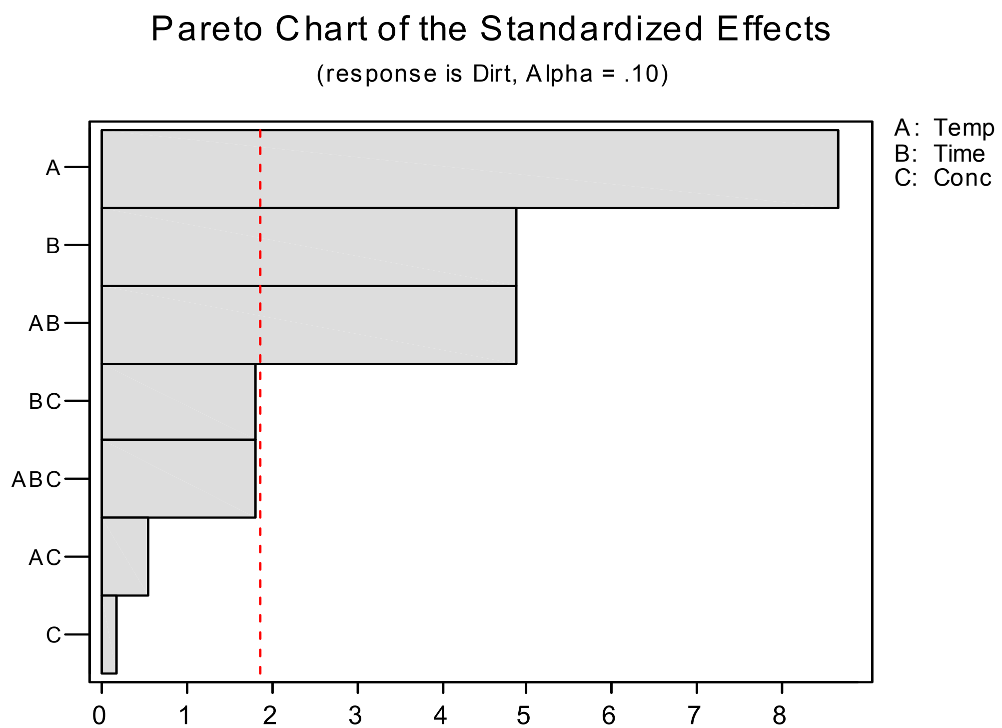
Interpreting Results

➢ STAT > DOE > FACTORIAL PLOTS > Interaction plot



Interpreting Results

➢ STAT > DOE > ANALYZE FACTORIAL DESIGN > GRAPHS > Pareto



➢ Pareto chart shows both magnitude & importance of an effect. Any effect extending past the reference line is significant

Beverages Industry Example

Consider that xyz is interested in obtaining more uniform fill heights in the bottles. The filling machine theoretically fills each bottle to the correct target height, but in practice, there is variation around this target, and the bottler would like to understand better the sources of this variability and eventually reduce it. There are three control factors

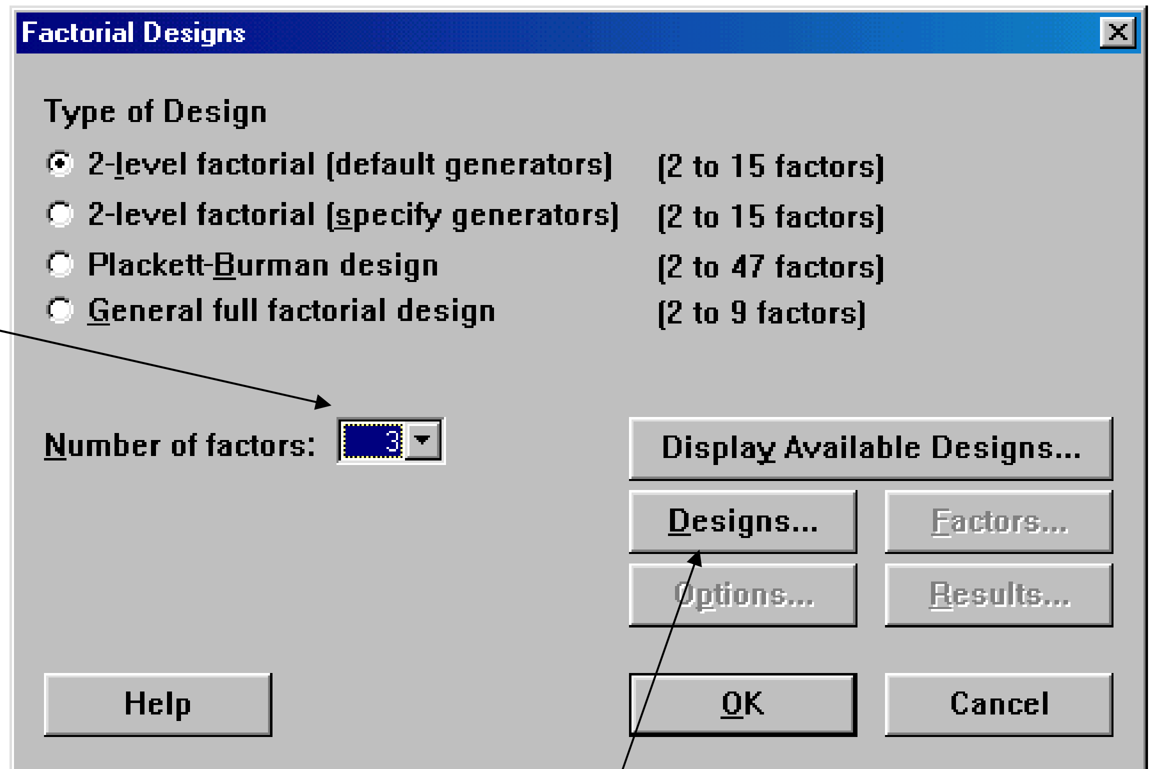
Factor	Unit	Level 1	Level 2
Carbonation	%	10	12
Operating Pressure	psi	25	30
Line Speed	BPM	600	650

The response is the average deviation from the target fill height observed in a production run of bottles.

Creating Designs in Minitab

- STAT > DOE > CREATE FACTORIAL DESIGN

3 factors relating to
Temperature,
Time &
Concentration



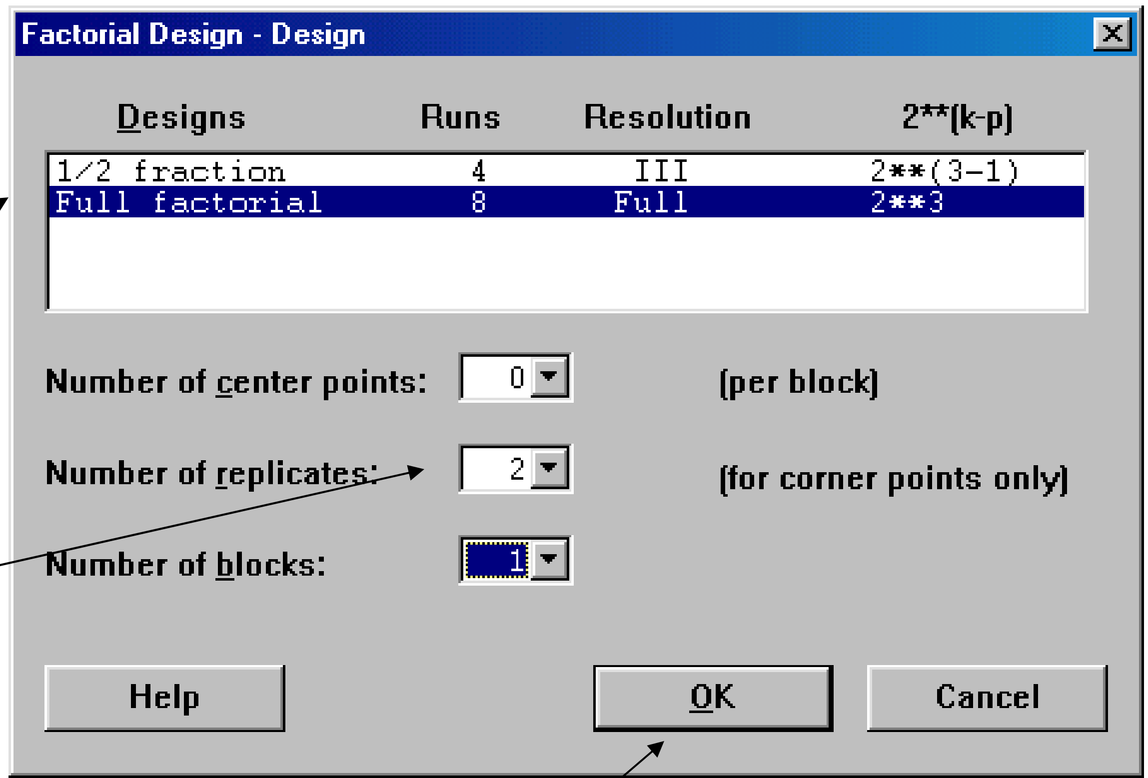
Click on Designs

Creating Designs in Minitab

- STAT > DOE > CREATE FACTORIAL DESIGN > DESIGNS

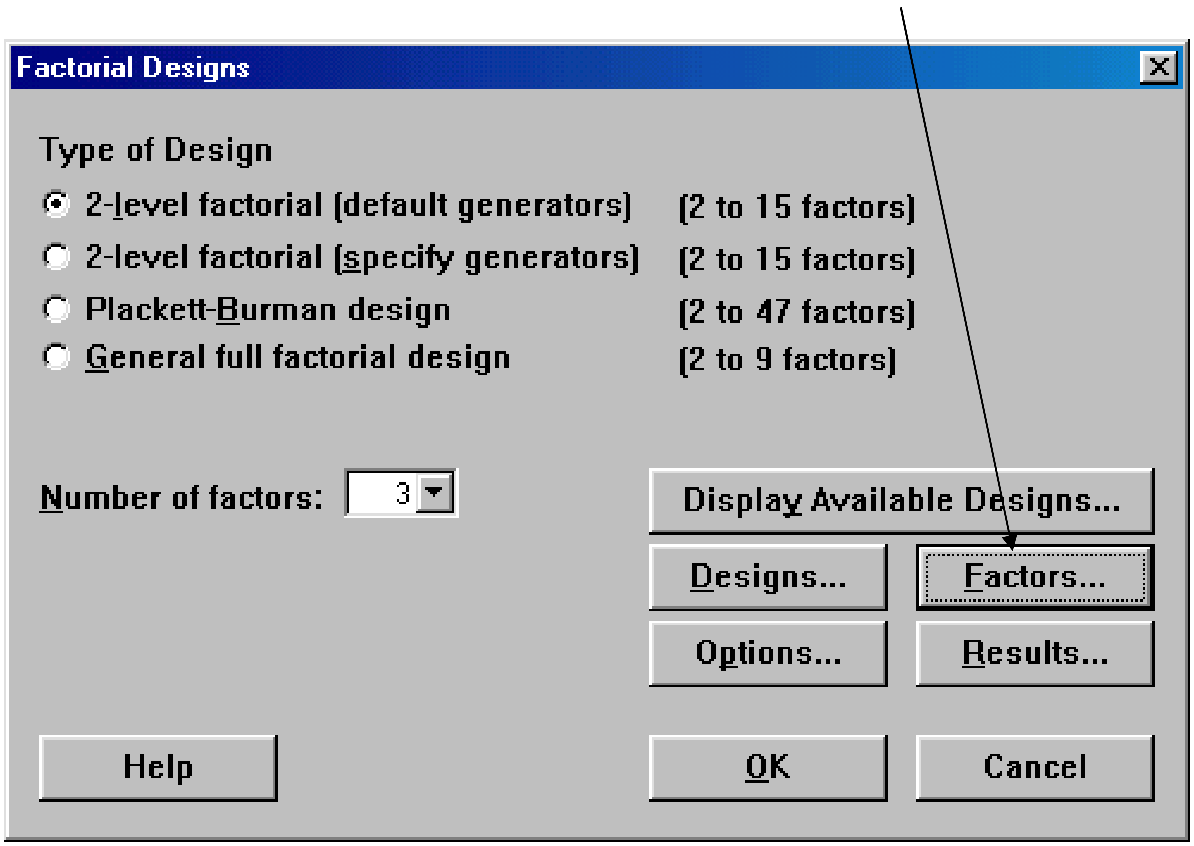
Choose a full factorial or fractional factorial design

Choose two replicates



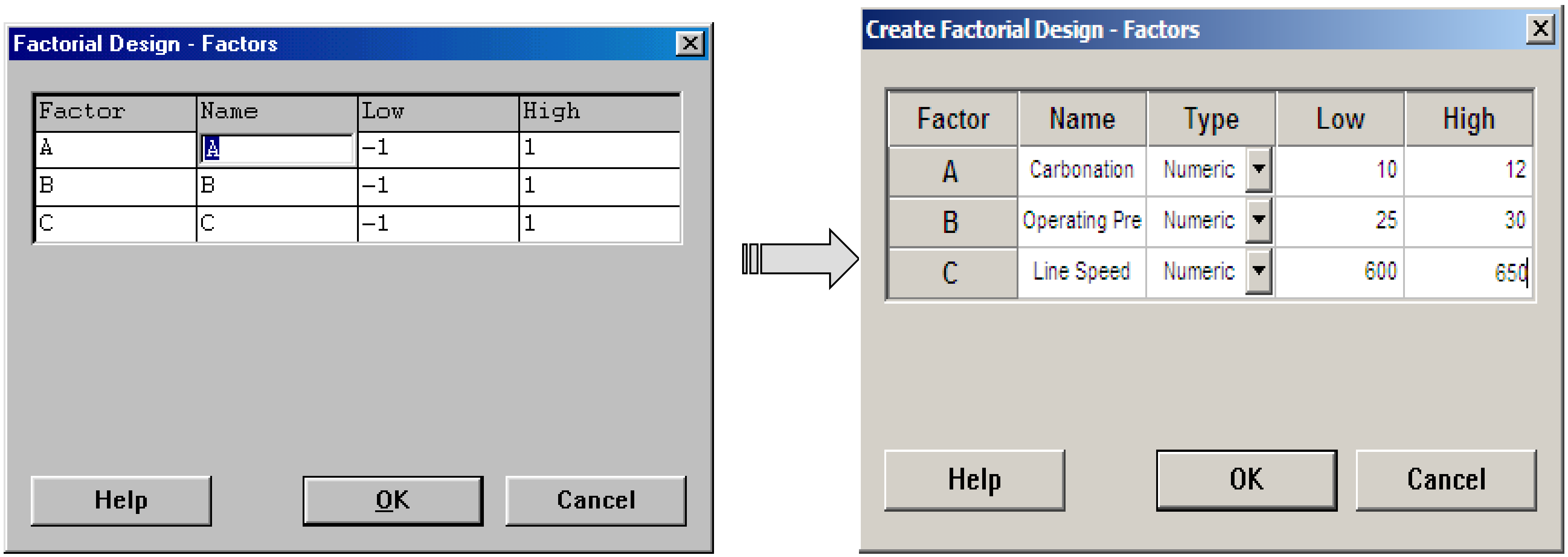
Creating Designs in Minitab

- STAT > DOE > CREATE FACTORIAL DESIGN > Click on factors



Creating Designs in Minitab

➢ STAT > DOE > CREATE FACTORIAL DESIGN > FACTORS



Click n OK

Creating Designs in Minitab

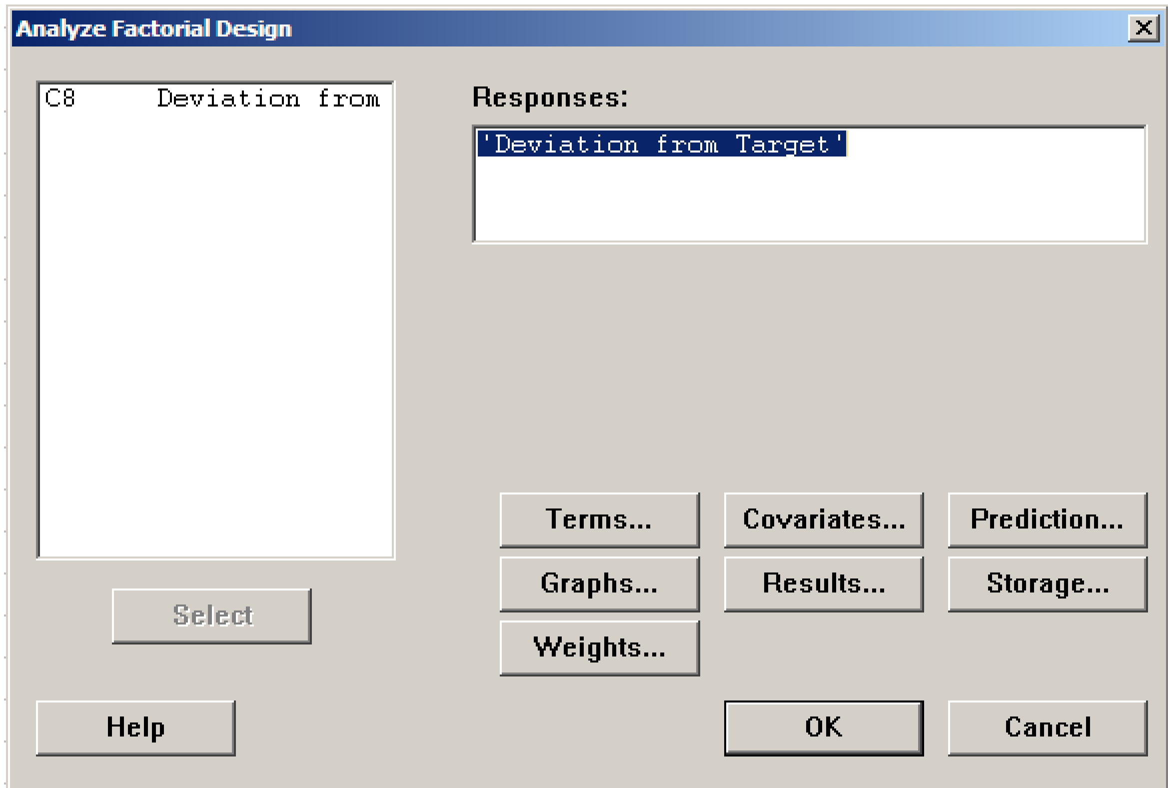
MINITAB - MINITAB.MPJ(DOE).MPJ - [Worksheet 1 ***]

File Edit Data Calc Stat Graph Editor Tools Window Help

↓	C1	C2	C3	C4	C5	C6	C7	C8
	StdOrder	RunOrder	CenterPt	Blocks	Carbonation	Operating Pressure	Line Speed	Deviation from Target
1	7	1	1	1	10	30	650	5
2	6	2	1	1	12	25	650	-2
3	5	3	1	1	10	25	650	1
4	4	4	1	1	12	30	600	3
5	13	5	1	1	10	25	650	-5
6	11	6	1	1	10	30	600	3
7	8	7	1	1	12	30	650	-4
8	1	8	1	1	10	25	600	-1
9	12	9	1	1	12	30	600	0
10	2	10	1	1	12	25	600	6
11	10	11	1	1	12	25	600	-1
12	16	12	1	1	12	30	650	-3
13	9	13	1	1	10	25	600	-9
14	3	14	1	1	10	30	600	1
15	14	15	1	1	12	25	650	-4
16	15	16	1	1	10	30	650	-2

Creating Designs in Minitab

- STAT > DOE > ANALYZE FACTORIAL DESIGN



- 'Analyze factorial design' will be enabled only if Minitab was used to create the design

Following is the Minitab Output:

Estimated Effects and Coefficients for Deviation from Target (coded units)

Term	Effect	Coef	SE Coef	T	P
Constant		-0.750	0.9186	-0.82	0.438
Carbonation	0.250	0.125	0.9186	0.14	0.895
Operating Pressure	2.250	1.125	0.9186	1.22	0.256
Line Speed	-2.000	-1.000	0.9186	-1.09	0.308
Carbonation*Operating Pressure	-3.000	-1.500	0.9186	-1.63	0.141
Carbonation*Line Speed	-3.250	-1.625	0.9186	-1.77	0.115
Operating Pressure*Line Speed	-0.750	-0.375	0.9186	-0.41	0.694
Carbonation*Operating Pressure* Line Speed	1.000	0.500	0.9186	0.54	0.601

S = 3.67423   R-Sq = 52.84%   R-Sq(adj) = 11.57%

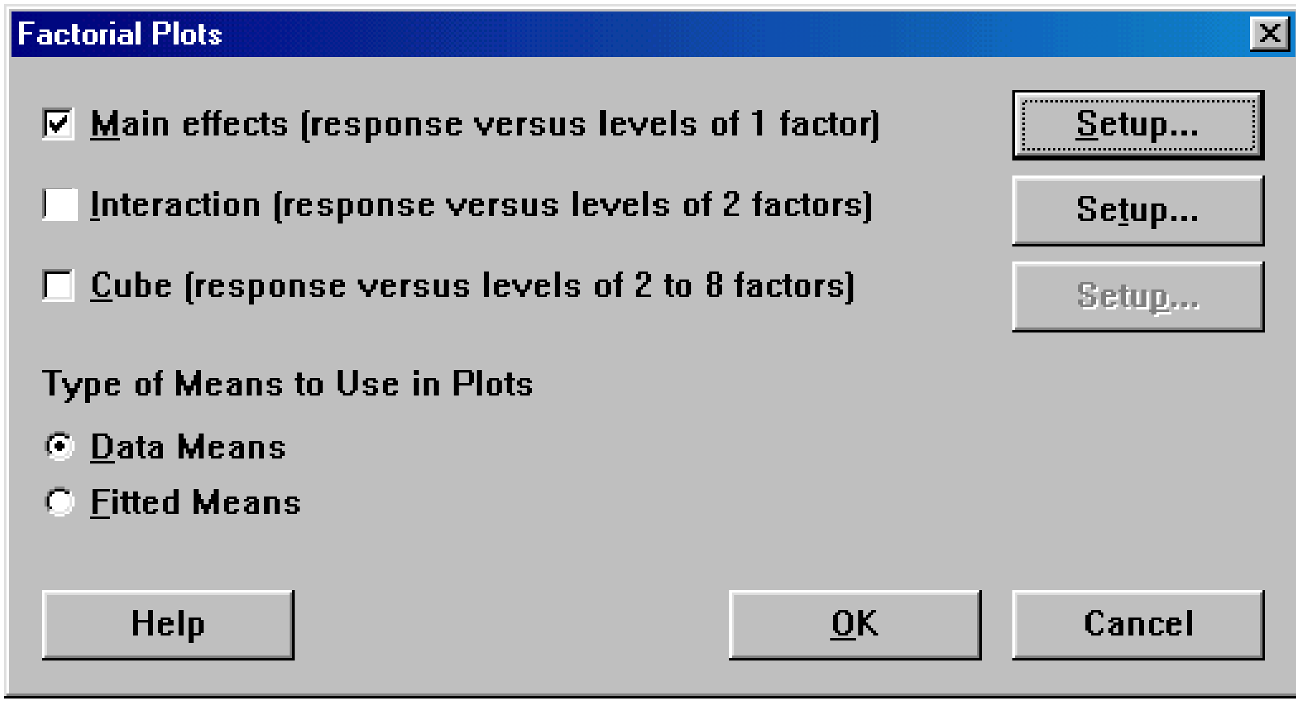
Analysis of Variance for Deviation from Target (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	36.500	36.500	12.167	0.90	0.482
2-Way Interactions	3	80.500	80.500	26.833	1.99	0.194
3-Way Interactions	1	4.000	4.000	4.000	0.30	0.601
Residual Error	8	108.000	108.000	13.500		
Pure Error	8	108.000	108.000	13.500		
Total	15	229.000				

Interpreting Results

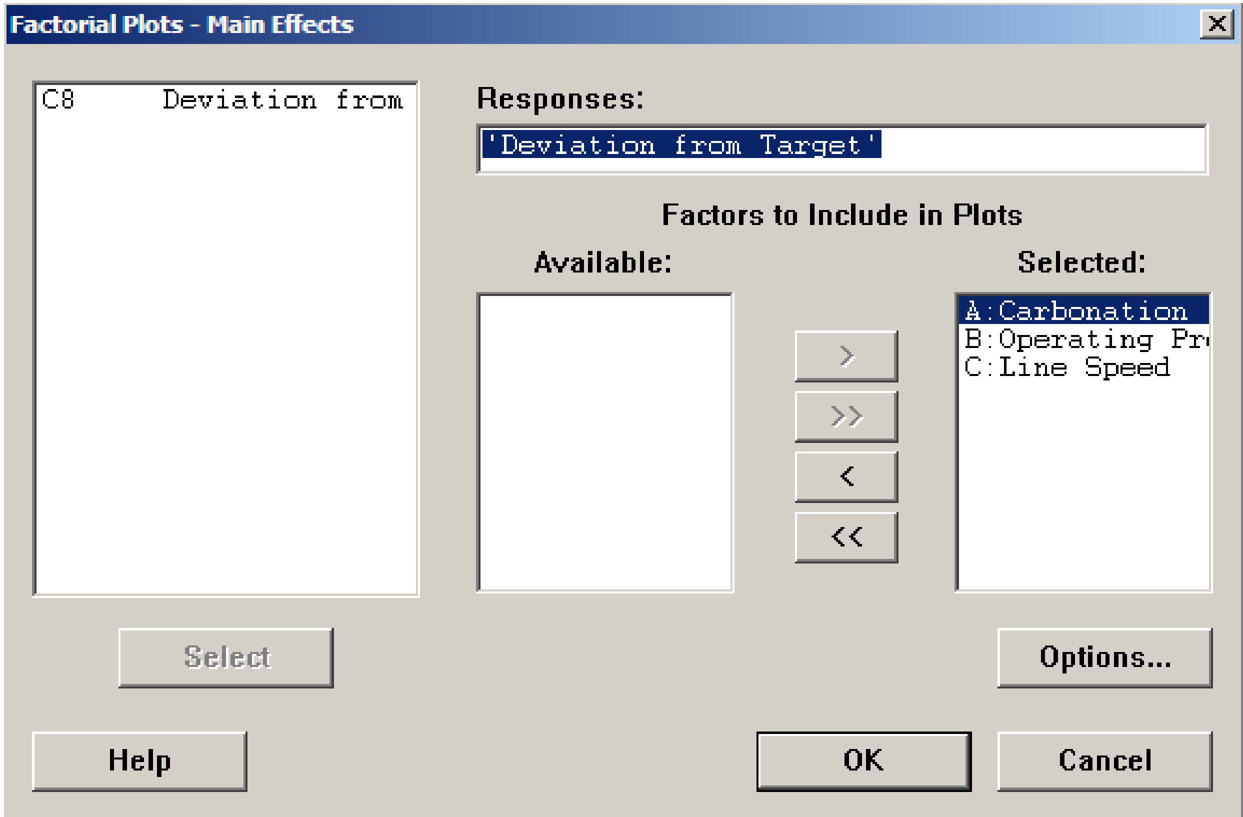
- STAT > DOE > FACTORIAL PLOTS

Click on Setup

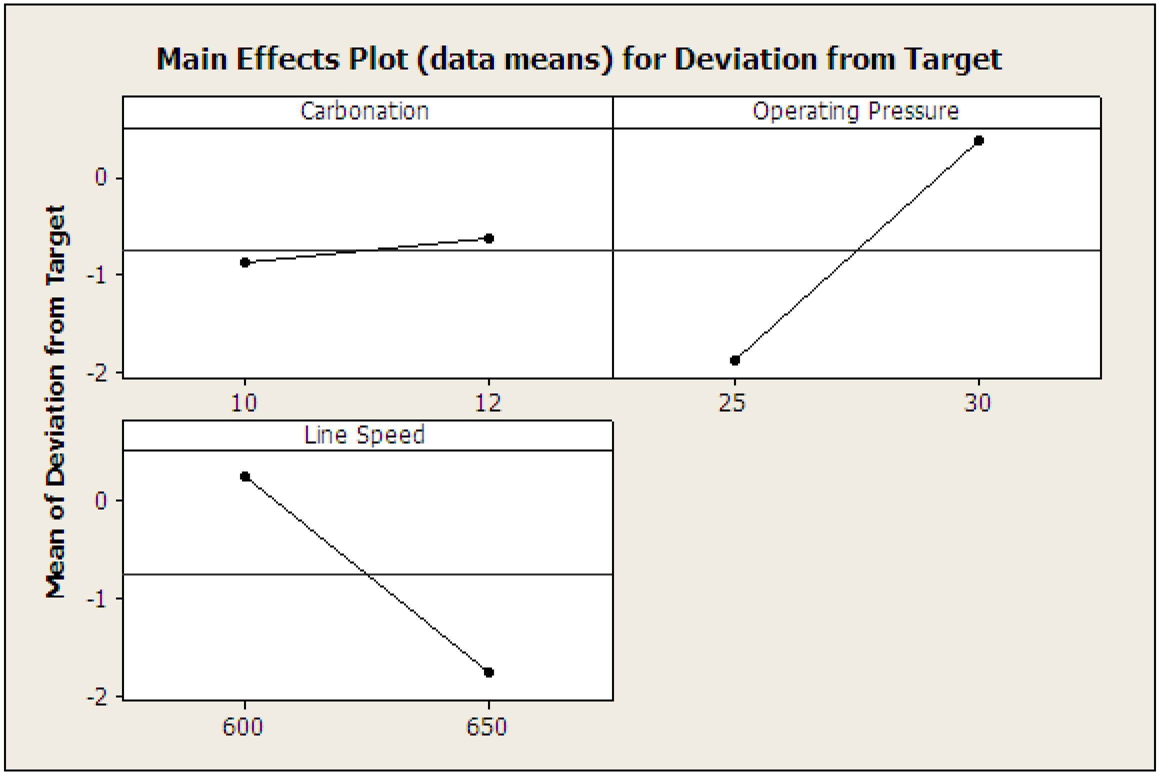


Interpreting Results

- STAT > DOE > FACTORIAL PLOTS > Main effects plot



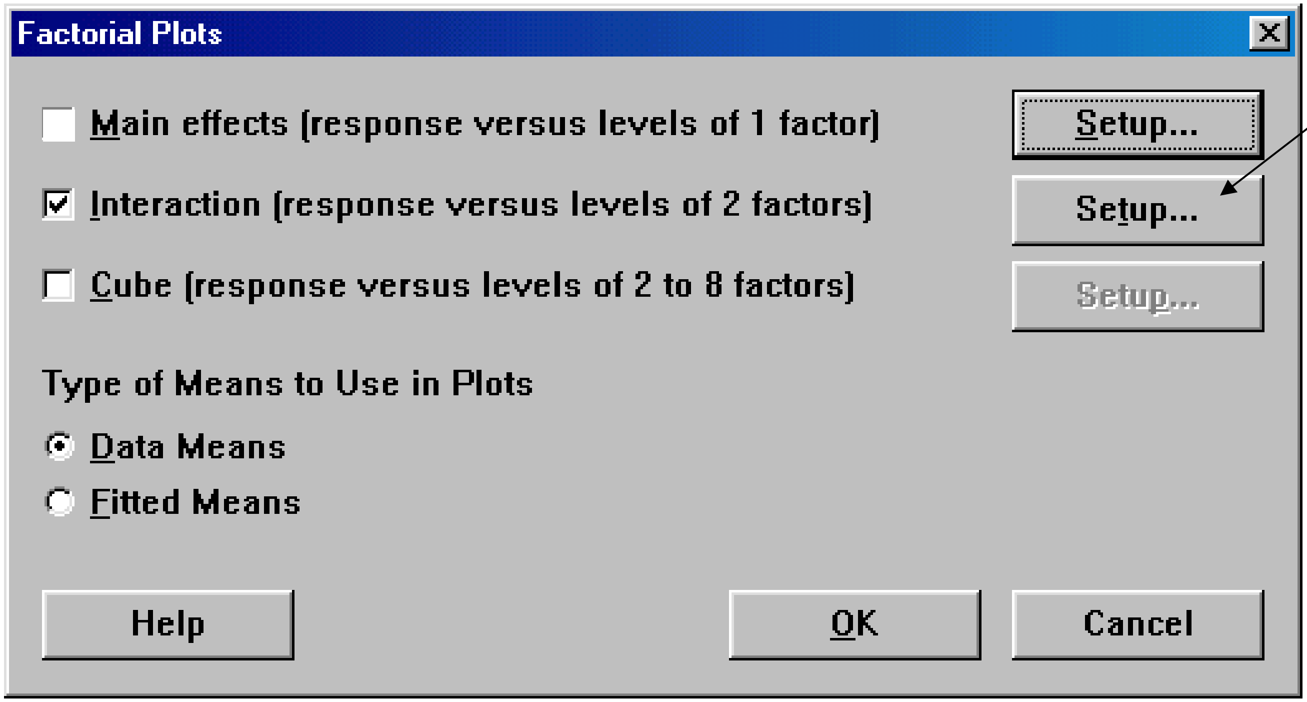
Interpreting Results



Interpreting Results

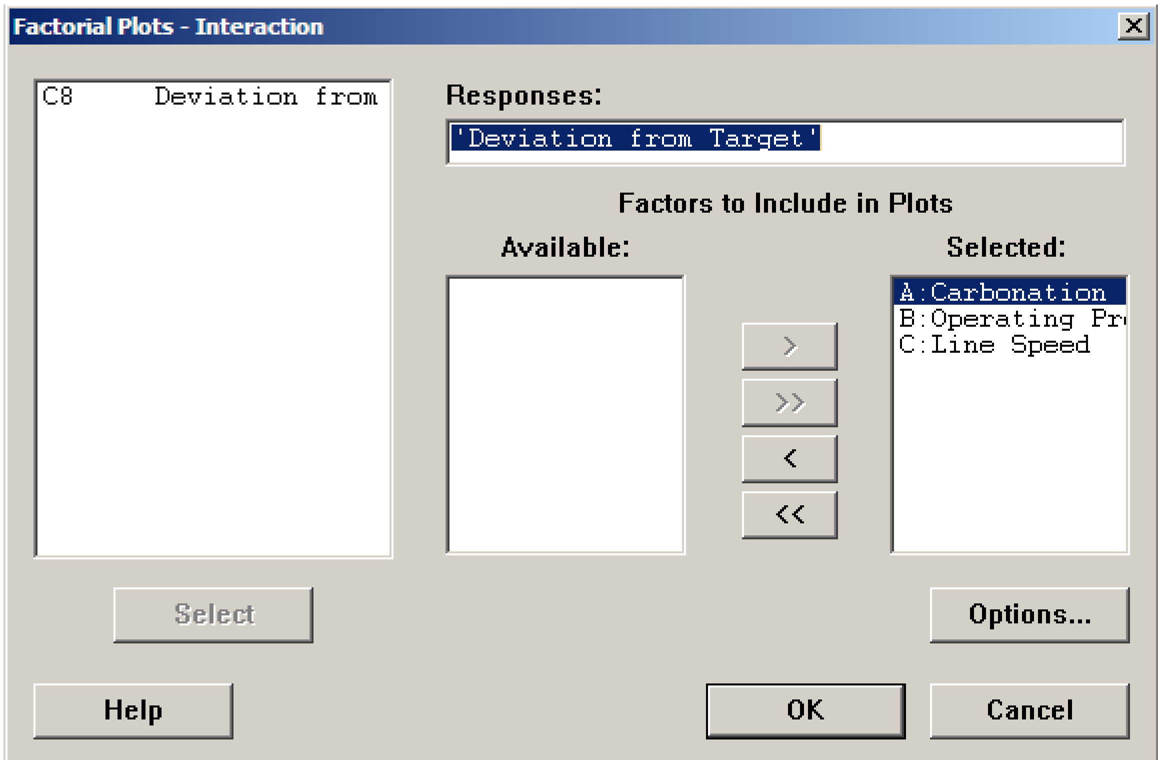
➢ STAT > DOE > FACTORIAL PLOTS

Click on Setup



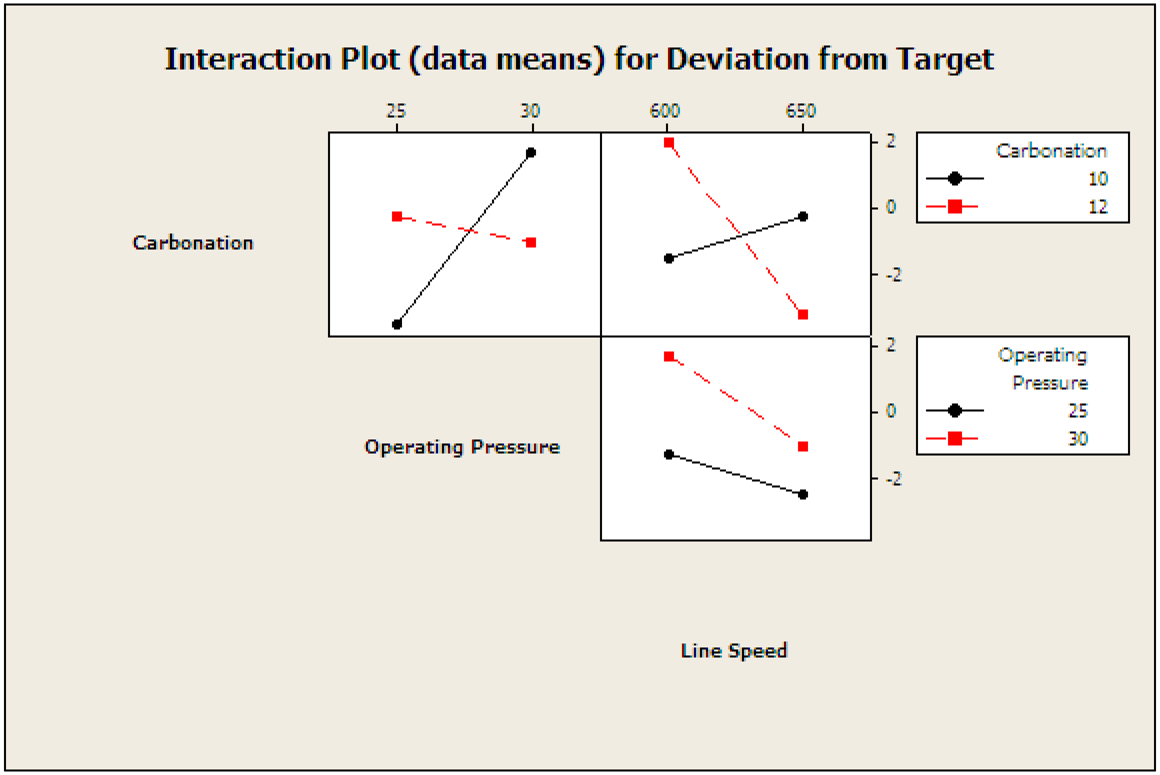
Interpreting Results

- STAT > DOE > FACTORIAL PLOTS > Interaction plot



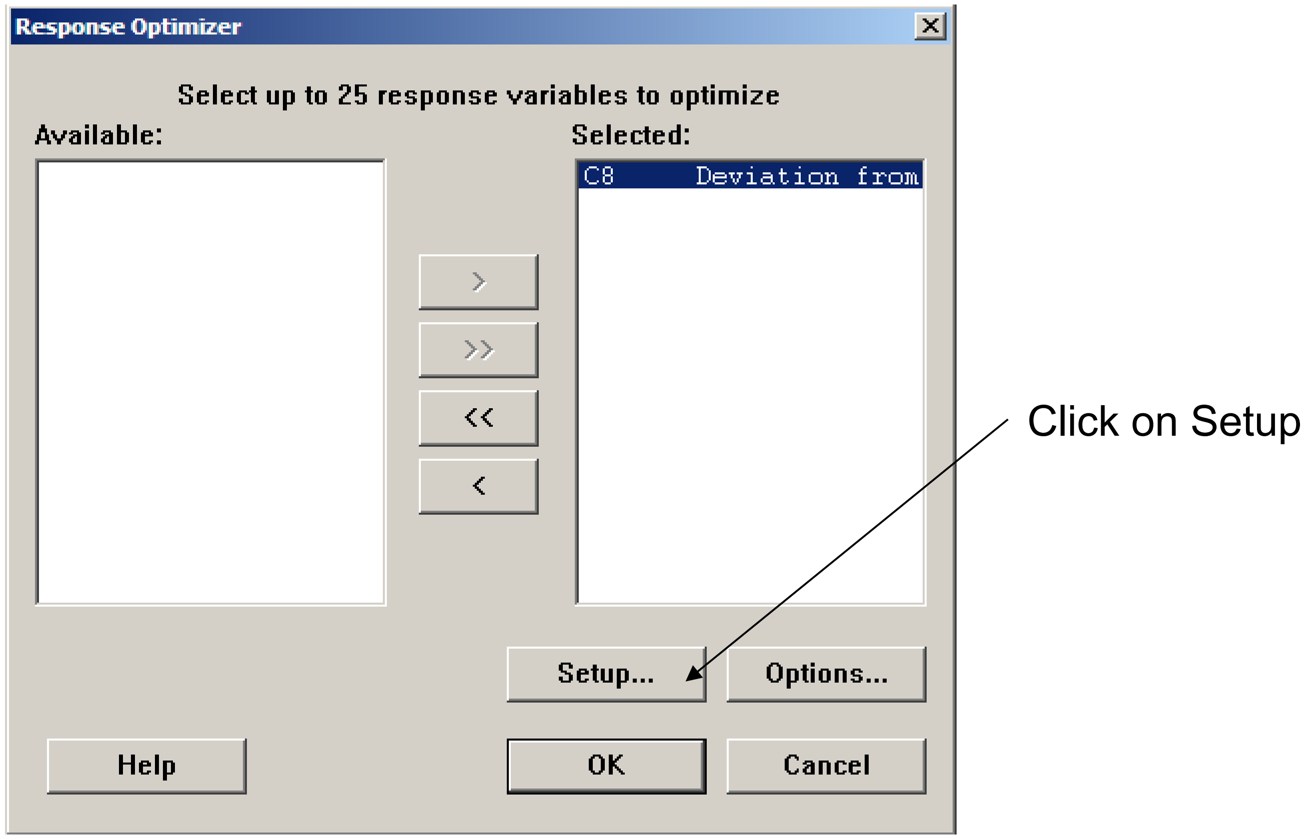
Interpreting Results

- STAT > DOE > FACTORIAL PLOTS > Interaction plot



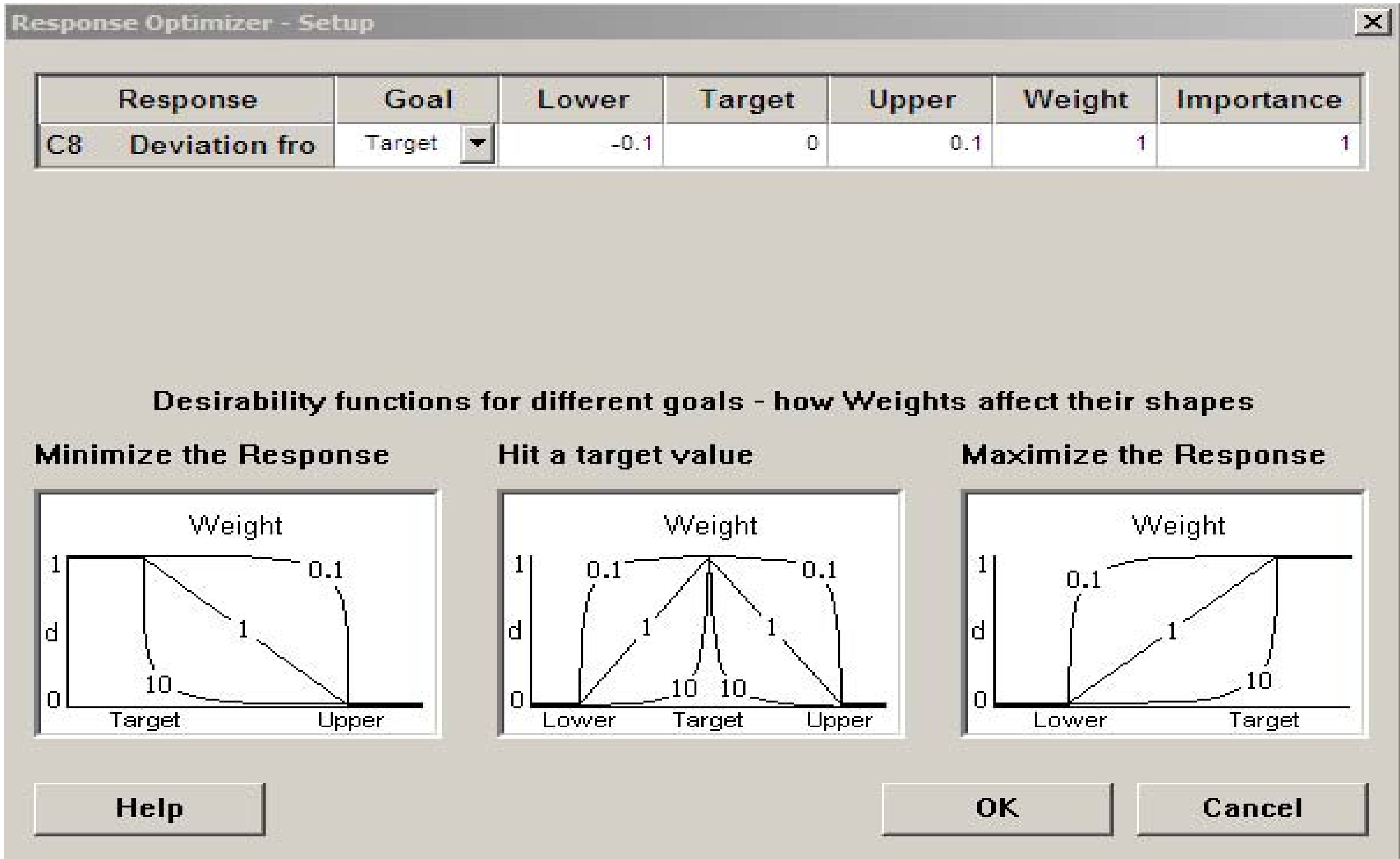
Response Optimization (Min Deviation from target)

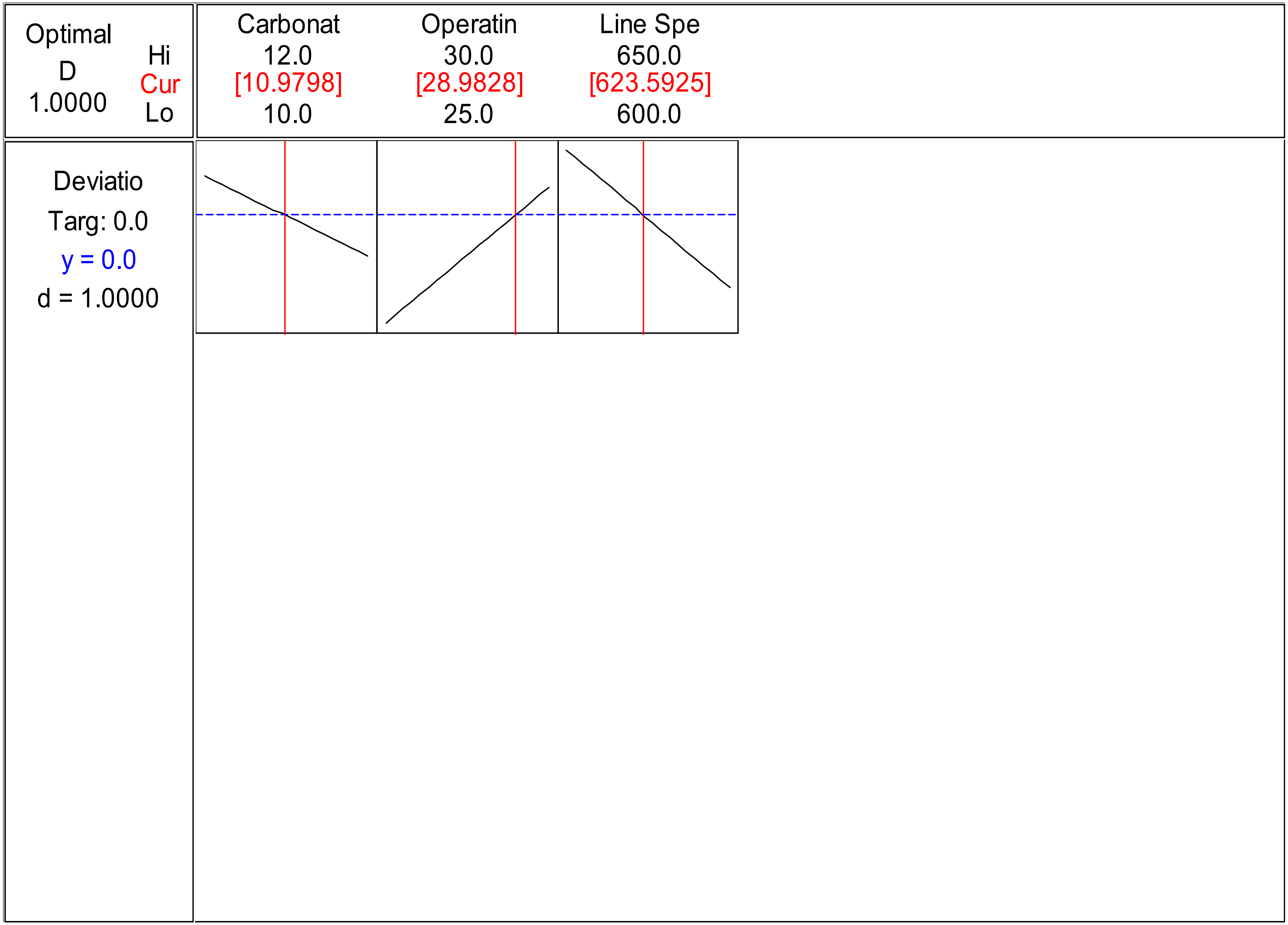
➢ STAT > DOE > FACTORIAL > Response optimizer



Response Optimization (Min Deviation from target)

Enter Goal = Target
Lower = -0.1
Upper = 0.1




Optimal
D
1.0000
Hi
Cur
Lo
Carbonat
12.0
[10.9798]
10.0
Operatin
30.0
[28.9828]
25.0
Line Spe
650.0
[623.5925]
600.0
Deviatio
Targ: 0.0
y = 0.0
d = 1.0000

Fractional Factorial Experiments

Why do Fractional Factorial Experiments?

- As the number of factors increases, so do the number of runs
  - 2x2 Factorial = 4 runs
  - 2x2x2 Factorial = 8 runs
  - 2x2x2x2 Factorial = 16 runs

  etc.
- If the experimenter can assume higher order interactions are negligible, it is possible to do a *fraction* of the full factorial and still get good estimates of low-order interactions
- The major use of Fractional Factorials is for screening variables
  - A relatively large number of Factors can be evaluated in a relatively small number of runs

Factorial Experiments

- Successful factorials are based on:
  - The Sparsity of Effects Principle
    - ✓ Systems are usually driven by Main Effects and Low-order interactions
  - The Projective Property
    - ✓ Fractional Factorials can represent full-factorials once some effects demonstrate weakness
  - Sequential Experimentation
    - ✓ Fractional Factorials can be combined into more powerful designs
    - ✓ Half-Fractions can be "folded over" into a full factorial
    - ✓ By eliminating uninteresting Input Variables, fractions can become full factorials

Half-Fraction

Recall that table below is the expanded representation of a 2^3 Factorial design

Factor D

A	B	C	AXB	AXC	BXC	AXBXC
-1	-1	-1	1	1	1	-1
1	-1	-1	-1	-1	1	1
-1	1	-1	-1	1	-1	1
1	1	-1	1	-1	-1	-1
-1	-1	1	1	-1	-1	1
1	-1	1	-1	1	-1	-1
-1	1	1	-1	-1	1	-1
1	1	1	1	1	1	1

Suppose we wanted to investigate four Input Variables but can not afford extra runs Since all the contrasts are independent (orthogonal) we can assign any interaction as the contrast to represent the fourth variable

Usually we select the highest order interaction and replace it with the additional factor. In this case, when we replace the AxBxC Interaction with Factor D, we say the ABC was aliased with D

Half-Fraction

The new design matrix looks like this:

A	B	C	D
-1	-1	-1	-1
1	-1	-1	1
-1	1	-1	1
1	1	-1	-1
-1	-1	1	1
1	-1	1	-1
-1	1	1	-1
1	1	1	1

This is a Half-fraction of a 2^4 design

Instead of 16 runs, we only need 8 runs to evaluate 4 factors

This is considered a Resolution IV design

Half Fraction

We would call this a half-fraction since a full 2^4 Factorial would take 16 runs to complete. Here we can estimate 4 factors in 8 runs.

But there is a cost: We lost the higher order interaction. When assessing what we have to lose, we use the concept of Resolution.

- *Resolution III Designs*
  - No main effects are aliased with other Main Effects
  - Main Effects aliased with two-factor interactions
- *Resolution IV Designs*
  - No Main Effect aliased with other Main Effects or with two-factor interactions
  - Two-factor interactions aliased with other two-factor interactions
- *Resolution V Designs*
  - Main Effects okay, Two-factor interactions aliased with 3-factor interactions

Notation

The general notation to designate a fractional factorial design is:

$$2_R^{k-p}$$

k is the number of factors to be investigated

2^{k-p} is the number of runs

R is the resolution

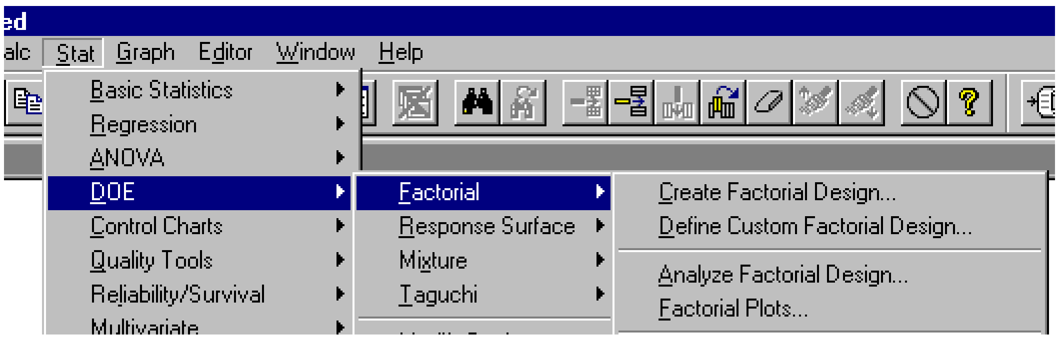
Example: The designation below means four factors will be investigated in $2^3 = 8$ runs. This design is resolution IV.

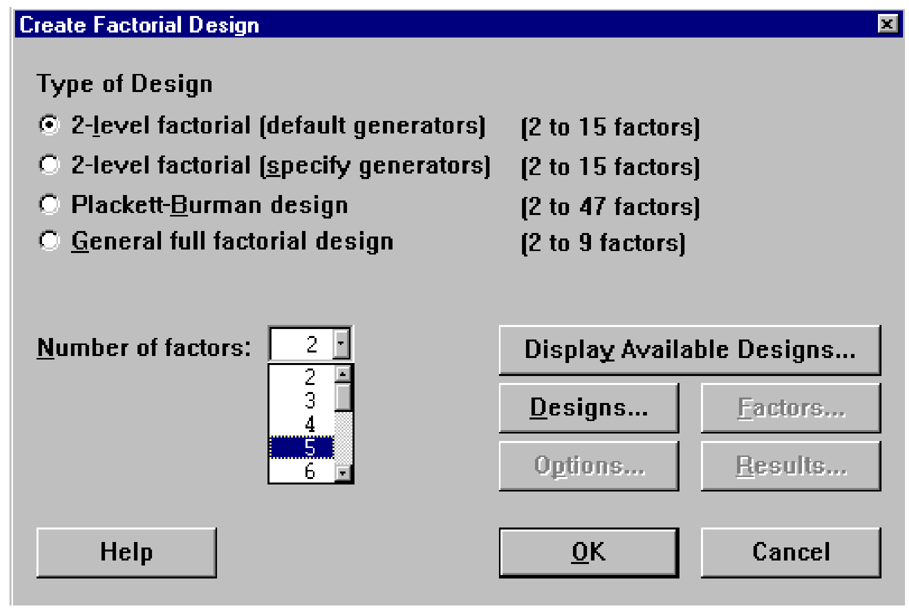
$$\boxed{2_{IV}^{4-1}}$$

$$\boxed{\frac{1}{2}2^5 = 2^{-1}2^5 = 2^5 2^{-1} = 2^{5-1}}$$

Fractional Factorials and Minitab

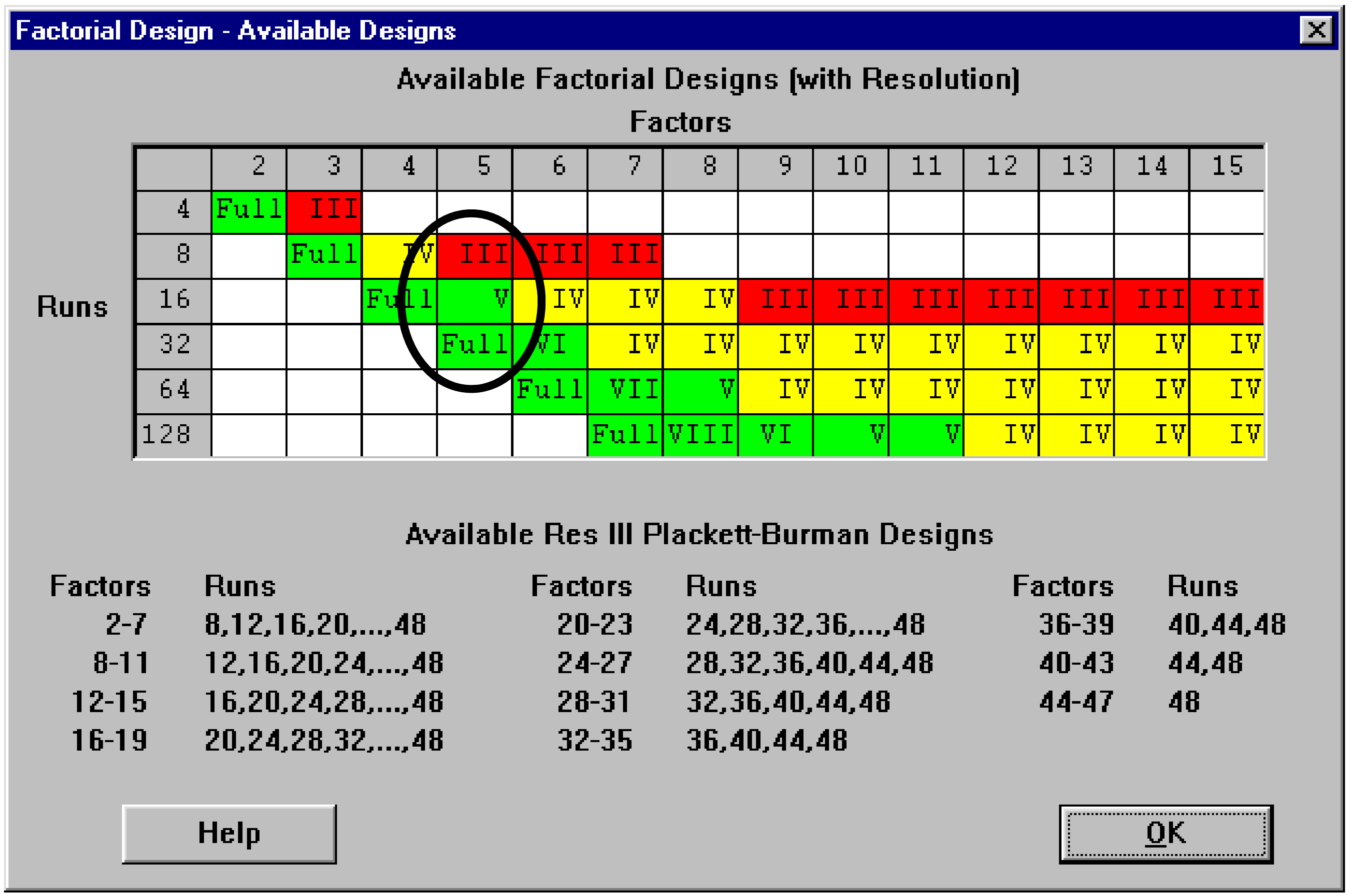
Let's take a look at the Minitab Dialog Boxes for designing a fractional factorial experiment with 5 factors





Fractional Factorials and Minitab

- Notice that for a 5 factor experiment we have two fractional factorial designs available
- Remember the aliasing for a Resolution III design



Factorial Design - Available Designs

Available Factorial Designs (with Resolution)

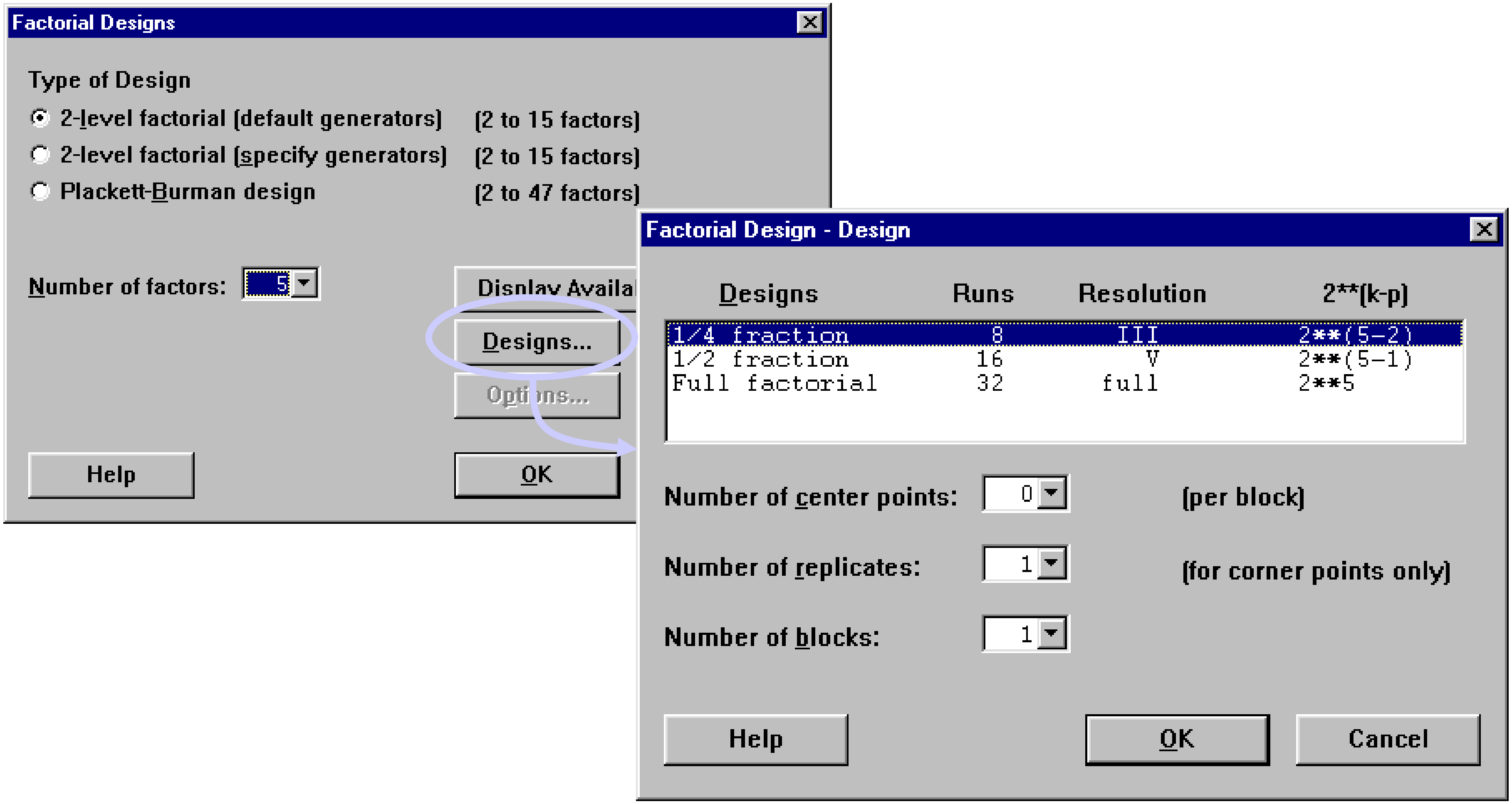
Factors

Runs	2	3	4	5	6	7	8	9	10	11	12	13	14	15
4	Full	III												
8		Full	IV	III	III	III								
16			Full	V	IV	IV	IV	III	III	III	III	III	III	III
32				Full	VI	IV	IV	IV	IV	IV	IV	IV	IV	IV
64					Full	VII	V	IV	IV	IV	IV	IV	IV	IV
128						Full	VIII	VI	V	V	IV	IV	IV	IV

Available Res III Plackett-Burman Designs

Factors	Runs	Factors	Runs	Factors	Runs
2-7	8,12,16,20,...,48	20-23	24,28,32,36,...,48	36-39	40,44,48
8-11	12,16,20,24,...,48	24-27	28,32,36,40,44,48	40-43	44,48
12-15	16,20,24,28,...,48	28-31	32,36,40,44,48	44-47	48
16-19	20,24,28,32,...,48	32-35	36,40,44,48		

Design Options



This table shows three options: Two fractional designs and the full factorial design

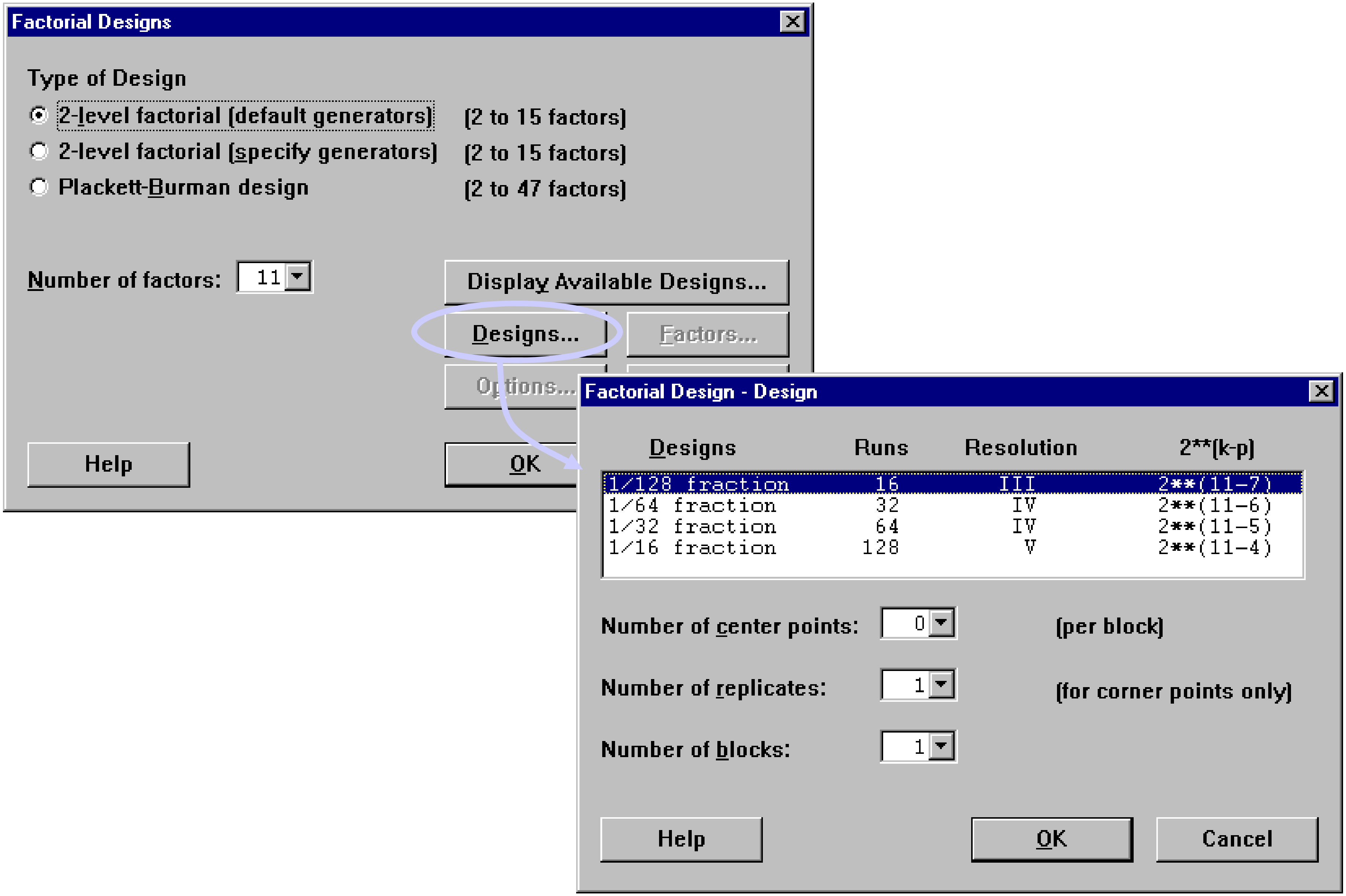
Exercise Continued

- Objective: To design and analyze a fractional factorial experiment using Minitab
- Procedure:
  - Use Minitab to setup a standard order Design Matrix
  - You only have funds to conduct 64 runs
- Choices:
  - Plan a 16 run fraction, with 4 replicates
  - Plan a 32 run, full factorial, with 2 replicates
  - Plan a 16 run, with 1 replicate, “keep the change”
    - Find an answer quicker
    - Save funds for follow up studies
    - Have a Pizza Party with the unspent funds!

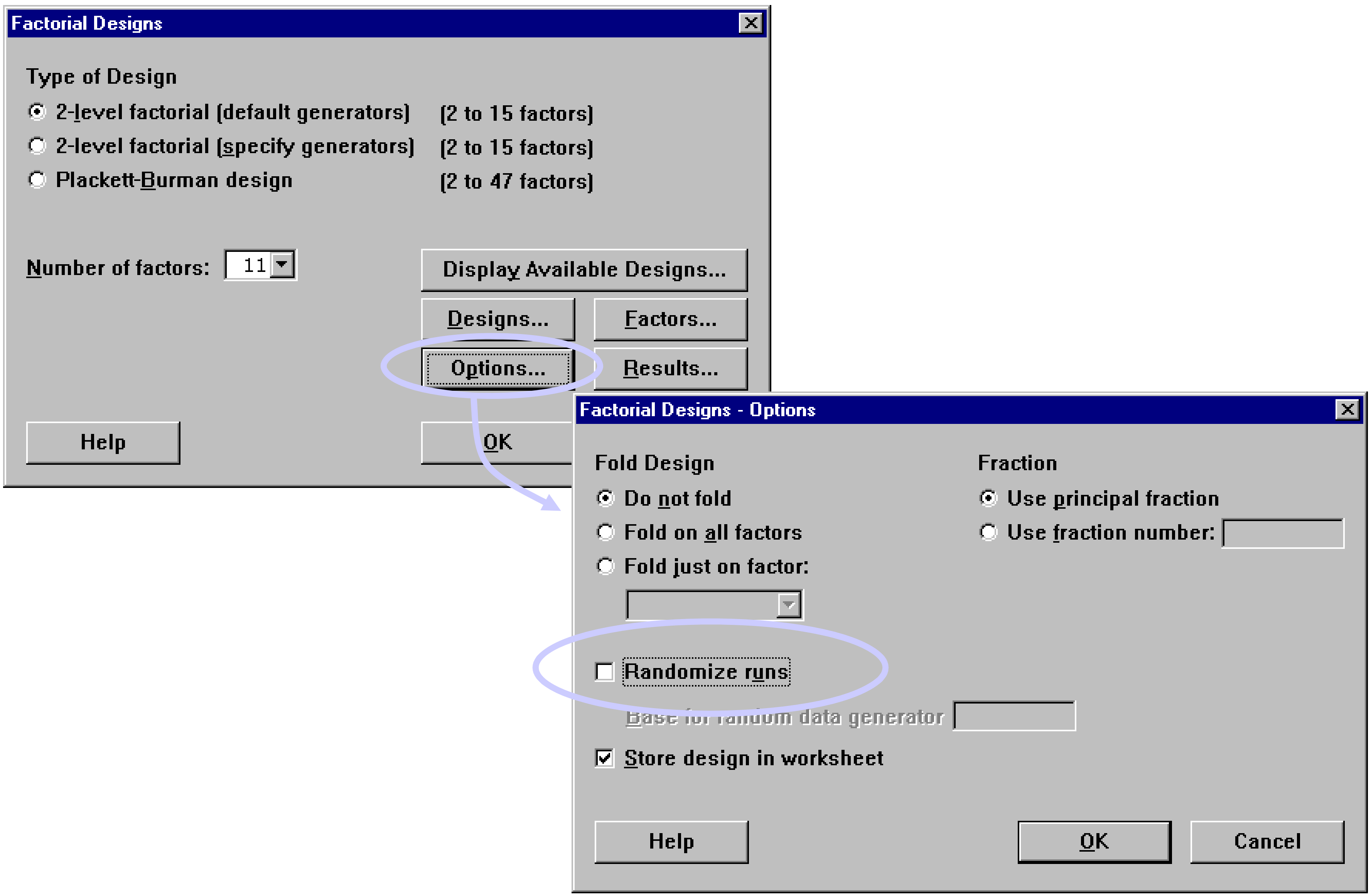
Minitab's Plan

StdOrder	RunOrder	CenterPt	Blocks	A	B	C	D	E
1	1	1	1	-1	-1	-1	-1	1
16	2	1	1	1	1	1	1	1
10	3	1	1	1	-1	-1	1	1
11	4	1	1	-1	1	-1	1	1
8	5	1	1	1	1	1	-1	-1
13	6	1	1	-1	-1	1	1	1
12	7	1	1	1	1	-1	1	-1
4	8	1	1	1	1	-1	-1	1
5	9	1	1	-1	-1	1	-1	-1
9	10	1	1	-1	-1	-1	1	-1
3	11	1	1	-1	1	-1	-1	-1
2	12	1	1	1	-1	-1	-1	-1
14	13	1	1	1	-1	1	1	-1
15	14	1	1	-1	1	1	1	-1
7	15	1	1	-1	1	1	-1	1
6	16	1	1	1	-1	1	-1	1

What if you had to look at ELEVEN inputs?



Defining the DOE



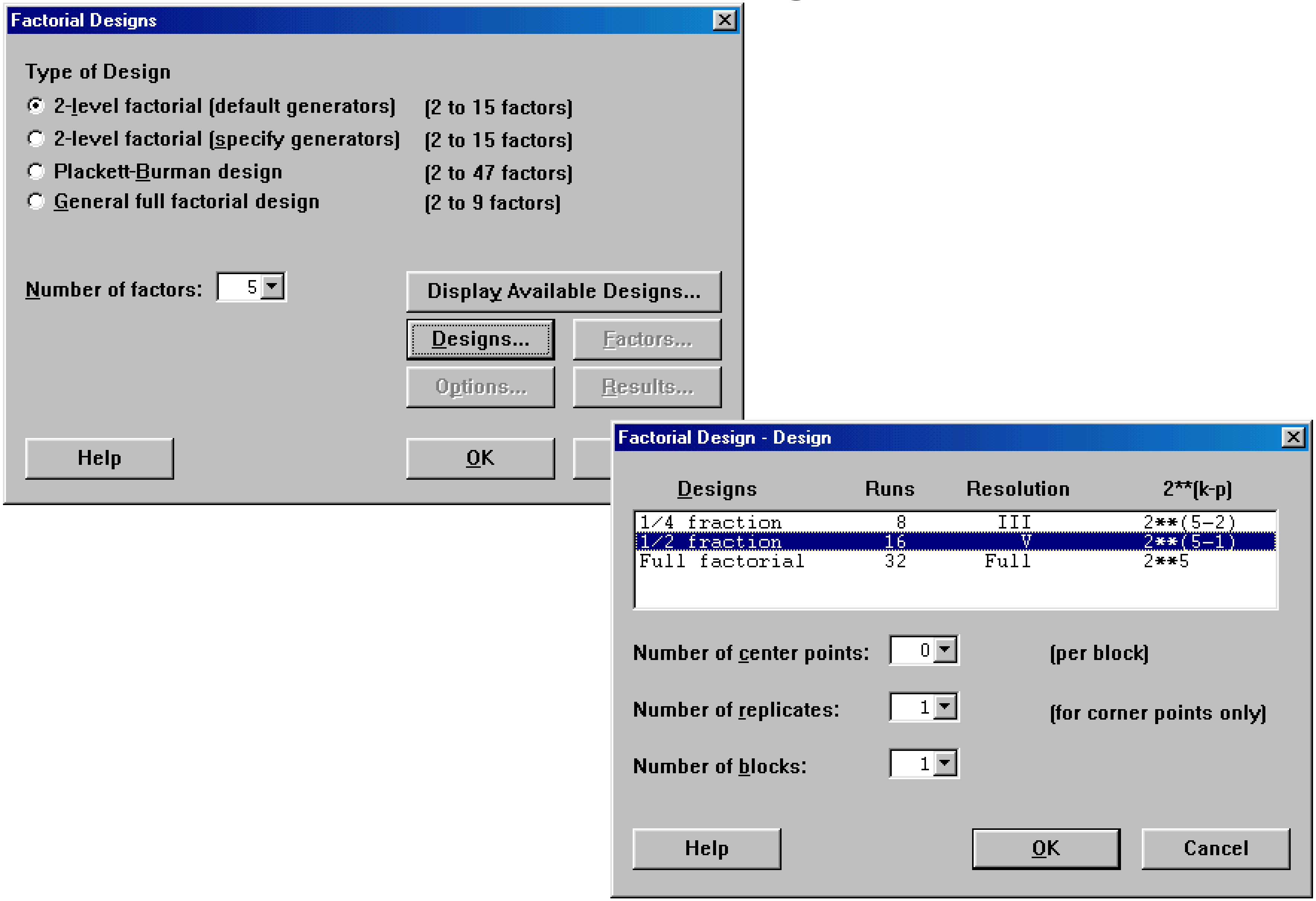
Result

A	B	C	D	E	F	G	H	J	K	L
1	-1	-1	-1	1	-1	1	1	-1	-1	-1
1	-1	1	-1	-1	1	-1	1	1	-1	1
-1	1	-1	-1	1	1	-1	1	-1	-1	1
-1	1	1	-1	-1	-1	1	1	1	-1	-1
1	-1	-1	1	1	1	-1	-1	1	-1	-1
-1	-1	1	1	1	-1	-1	1	1	1	-1
1	1	-1	1	-1	-1	-1	1	-1	1	-1
1	1	1	-1	1	-1	-1	-1	-1	1	1
-1	-1	1	-1	1	1	1	-1	-1	1	-1
1	1	-1	-1	-1	1	1	-1	1	1	-1
-1	1	-1	1	1	-1	1	-1	1	-1	1
-1	-1	-1	1	-1	1	1	1	-1	1	1
1	-1	1	1	-1	-1	1	-1	-1	-1	1
1	1	1	1	1	1	1	1	1	1	1
-1	-1	-1	-1	-1	-1	-1	-1	1	1	1
-1	1	1	1	-1	1	-1	-1	-1	-1	-1

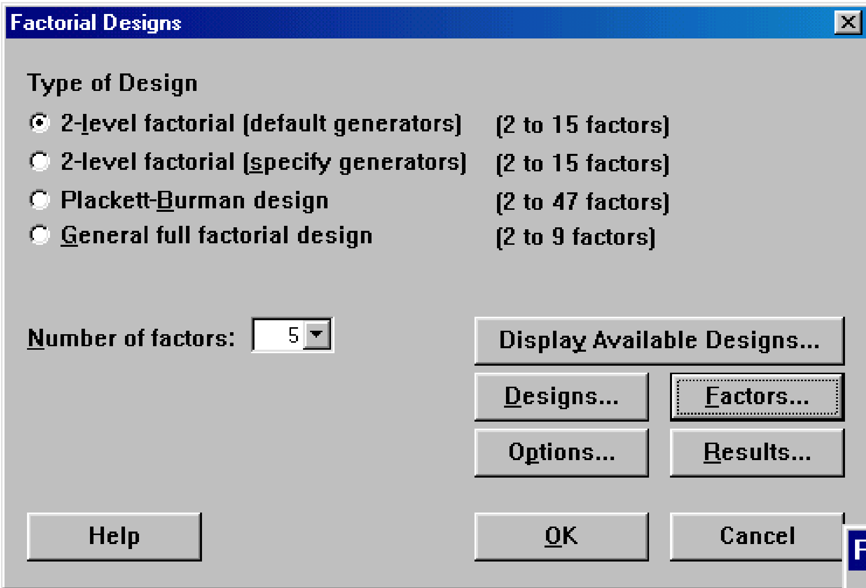
Exercise

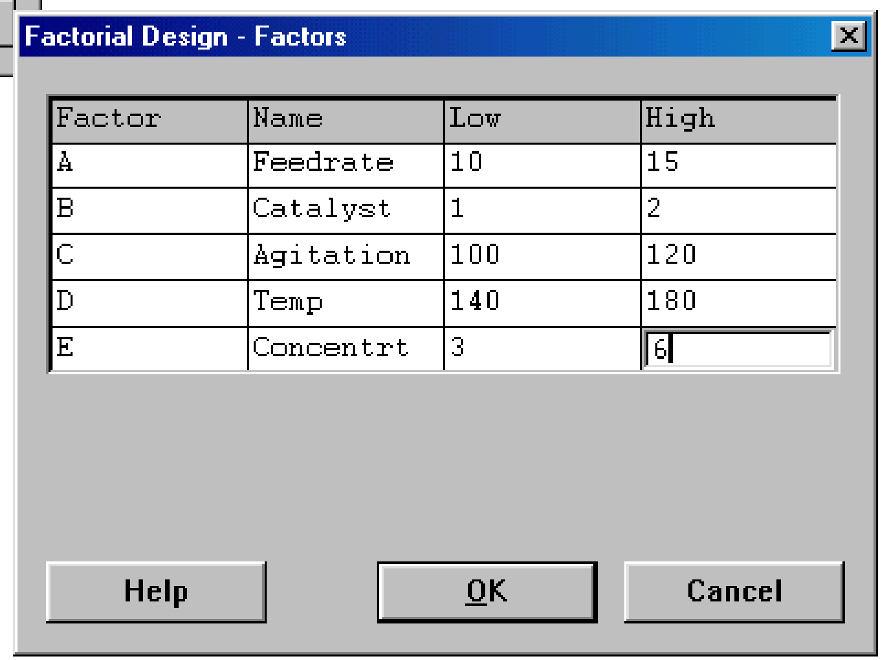
- Objective: To design and analyze a fractional factorial experiment using Minitab
- Output Variable: % Reacted
- Inputs:
  - Feed Rate (liters/minute) 10, 15
  - Catalyst (%) 1,2
  - Agitation Rate (rpm) 100, 120
  - Temperature (C) 140,180
  - Concentration (%) 3, 6
- Use Minitab to setup the Design Matrix
- You only have funds to do 16 runs
- Step 1: Name the columns for the Factors
- Step 2: Go to *Stat>DOE>Factorial>Create Factorial Design*

Exercise - Designs...

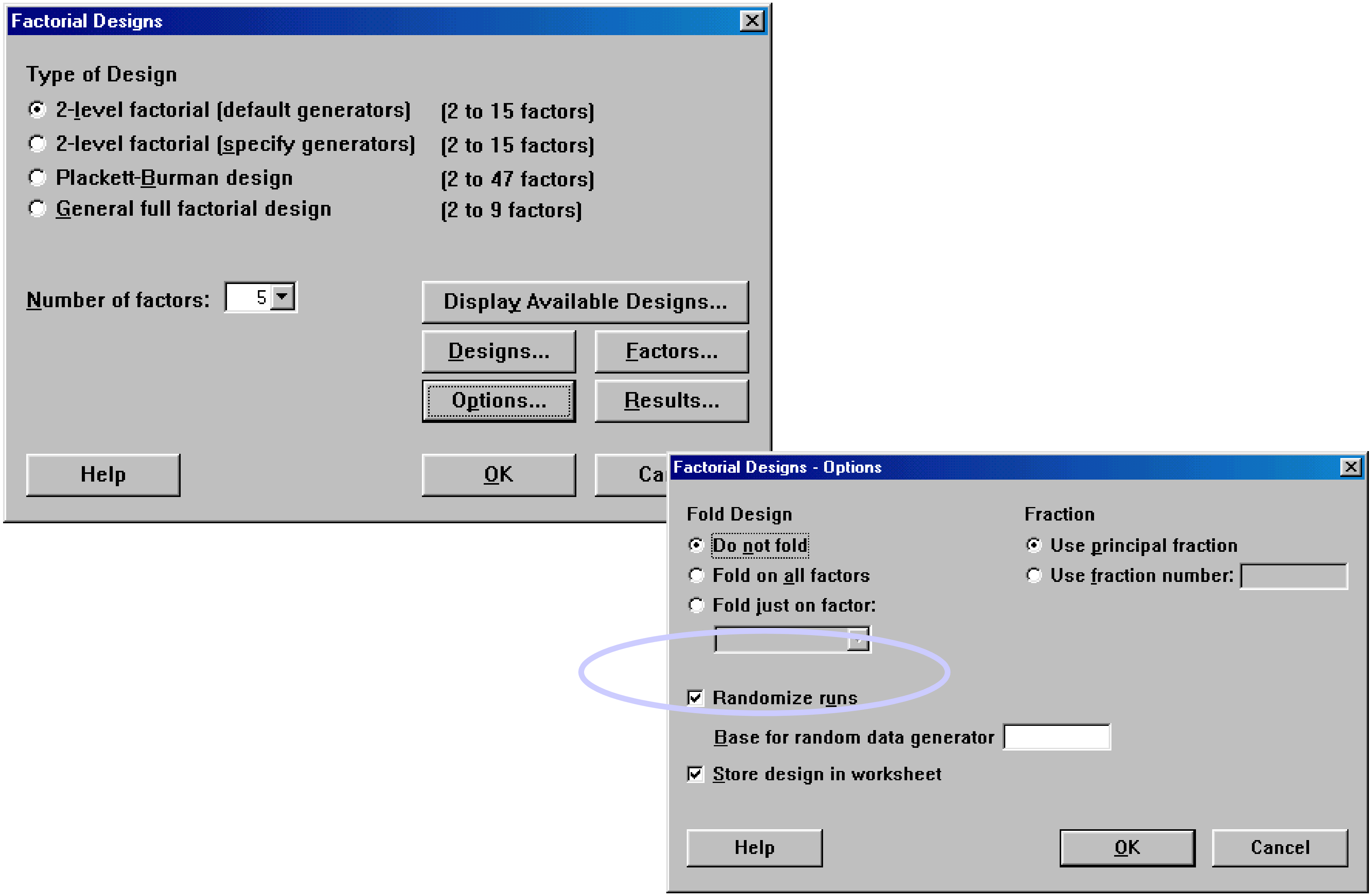


Exercise - Factors...





Exercise - Options...

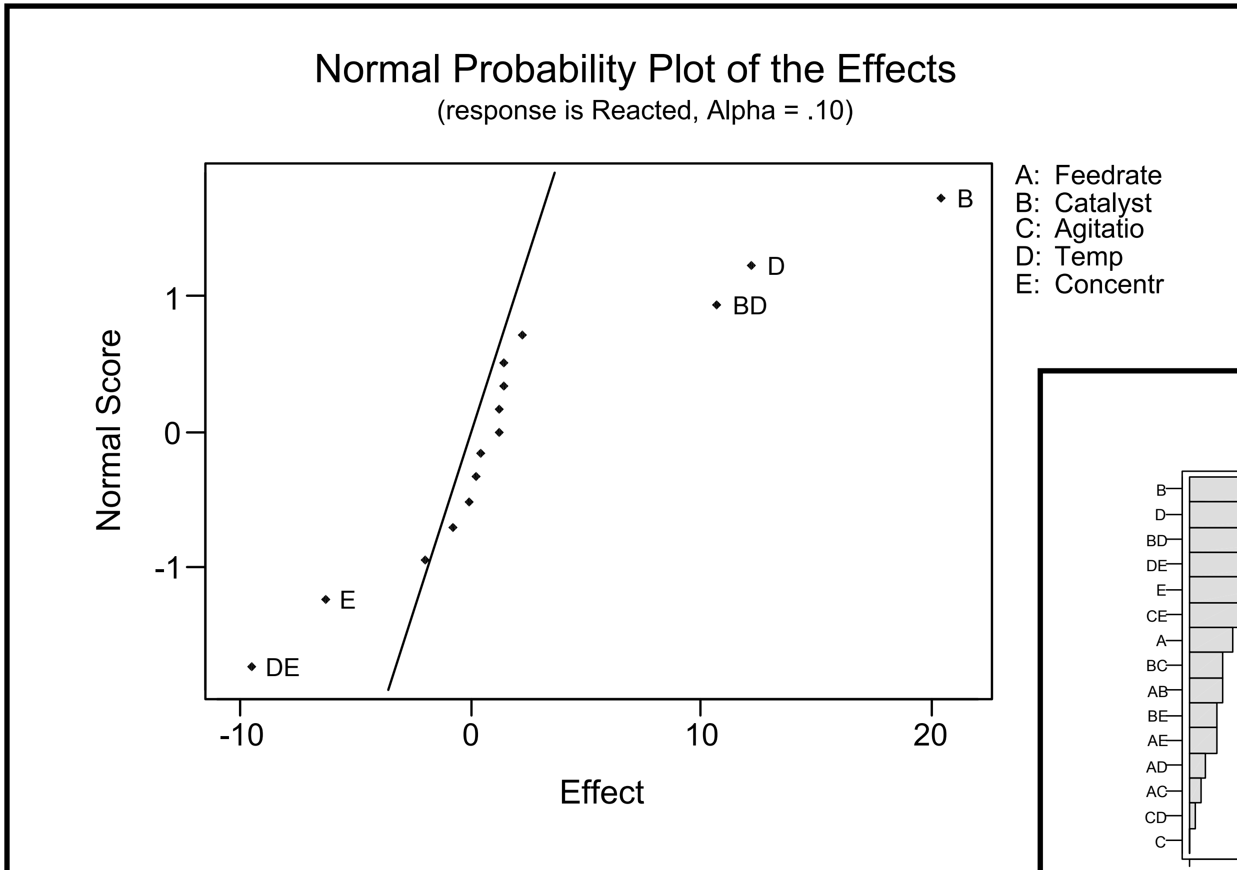


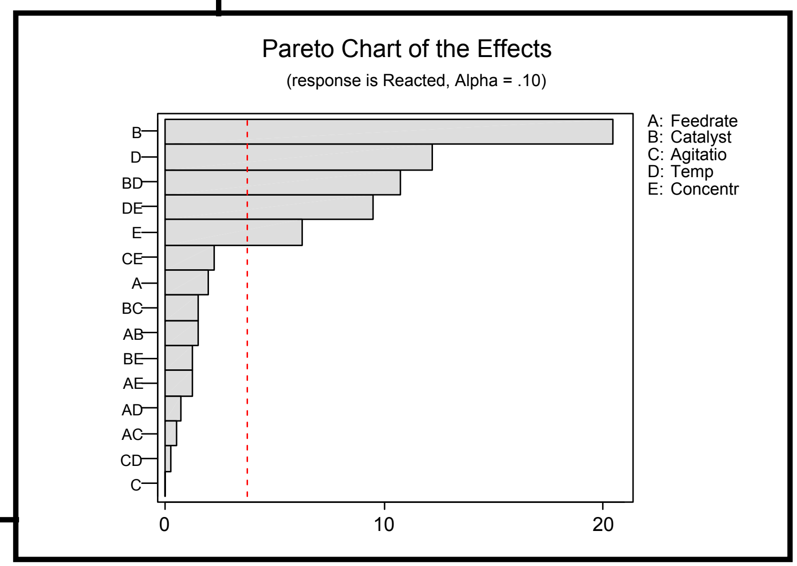
Design Matrix

	StdOrder	RunOrder	Blocks	Feedrate	Catalyst	Agitation	Temp	Concentrt
1	6		1	10	1	100	140	6
2	7		1	15	1	100	140	3
3	9		1	10	2	100	140	3
4	15		1	15	2	100	140	6
5	1		1	10	1	120	140	3
6	5		1	15	1	120	140	6
7	12		1	10	2	120	140	6
8	14		1	15	2	120	140	3
9	16		1	10	1	100	180	3
10	3		1	15	1	100	180	6
11	13		1	10	2	100	180	6
12	8		1	15	2	100	180	3
13	11		1	10	1	120	180	6
14	10		1	15	1	120	180	3
15	2		1	10	2	120	180	3
16	4		1	15	2	120	180	6

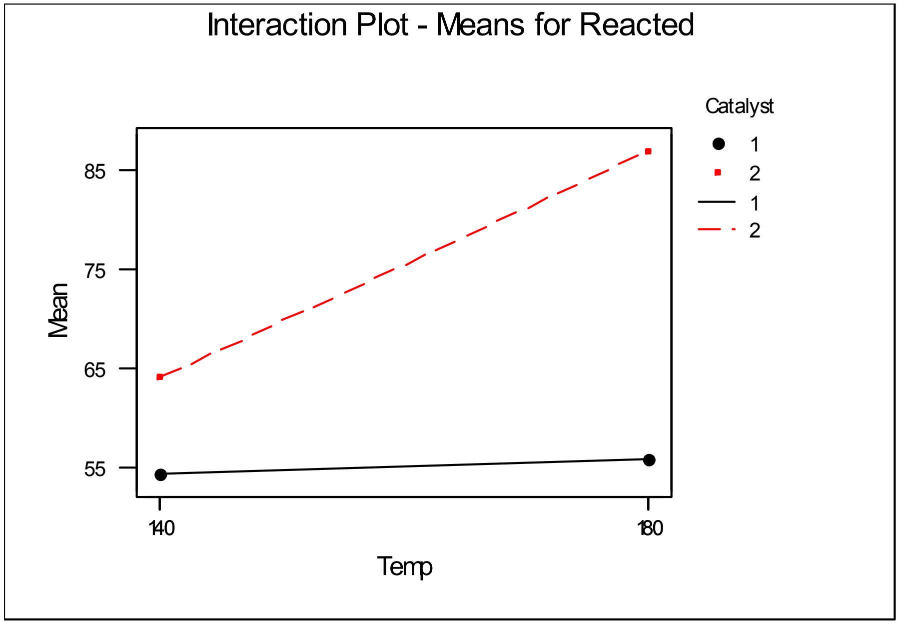
Exercise - Add Data

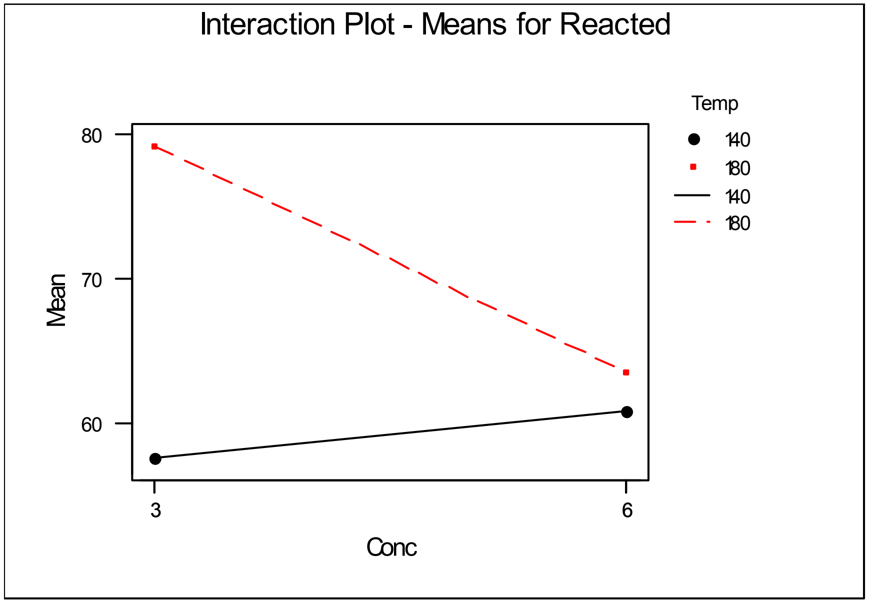
➢ Bring up Minitab File: BHH379.mtw and analyze the data





Interaction Plots





Implementation Plan

Critical Success Factors for Implementation

- **Critical Success Factors for solution implementation**
  - Buy in from sponsor
  - Buy in from key stakeholder
  - Strong communication plan that covers the following
    - ✓ Why this solution;
    - ✓ Expected benefits;
    - ✓ Time frame for testing or pilot;
    - ✓ Actionable after the test or pilot;
    - ✓ Responsibility

Cost Benefit Analysis

- What is Cost Benefit Analysis?
  - Financial benefits communication net of the costs incurred
  - Cost Benefit Analysis accelerates the buy in process and reassures bottom line benefit
- What are the benefit heads?
  - Revenue generation
  - Cost reduction
  - Improved quality
- And Costs could be incurred on
  - Training on the new solution
  - New equipment purchase for implementing the solution
  - Travel and living expenses incurred.

Implementation Plan

Implementation Plan		
Action	Responsibility	Date

RACI chart overview

Responsibility charting

The RACI technique has been designed to identify functional areas, key activities and provides management with decision points where ambiguities exist.

The approach enables management to actively participate in the process of systematically describing:

- activities
- decisions to be accomplished
- clarity of responsibilities

Guidelines

- Place accountability (A) and responsibility (R) at the level closest to the action
- There can only be one accountability (A) per activity
- Authority must accompany accountability
- Minimise the number of consults (C) and informs (I)
- All roles and responsibilities must be documented and communicated

RACI

A – accountable The buck stops here – yes/no authority

R – responsible The doer – working on the activity

C – consult In the loop – involved prior to decision/action

I – Inform Keep in the picture – needs to know of the decision/action

RACI - defined

Responsibility R	The individual(s) who actually completes the task, the doer. Responsibility can be shared. The degree of responsibility is determined by the individual with the "A"
Accountability A	The person who is ultimately responsible. Only one "A" can be assigned to a task
Consult C	The individual(s) to be consulted prior to a final decision or action. This incorporates two way communication
Inform I	The individual(s) who needs to be informed after a decision or action is taken, This is one way communication

Lean Six Sigma Green Belt

by
Rajiv Purkayastha
Six sigma MBB

Poka-yoke

What is Poke-yoke?

A method that uses sensor or other devices for catching errors that may pass by operators or assemblers.

Poka-yoke effects two key elements of ZDQ:

Identifying the defect immediately (Point of Origin Inspection)

Quick Feedback for Corrective Action

<u>How effective the system is depends on where it is used: Point of Origin or Informative Inspection.</u>

Poka-yoke detects an error, gives a warning, and can shuts down the process.

Ten Types of Human Mistakes

- Forgetfulness
- Misunderstanding
- Wrong identification
- Lack of experience
- Willful (ignoring rules or procedure)
- Inadvertent or sloppiness
- Slowness
- Lack of standardization
- Surprise (unexpected machine operation, etc.)
- Intentional (sabotage)

Poka-yoke Systems Govern the Process

Two Poka-Yoke System approaches are utilized lead to successful ZDQ:

1. **Control Approach**

   Shuts down the process when an error occurs.

   Keeps the “suspect” part in place when an operation is incomplete.

2. **Warning Approach**

   Signals the operator to stop the process and correct the problem.

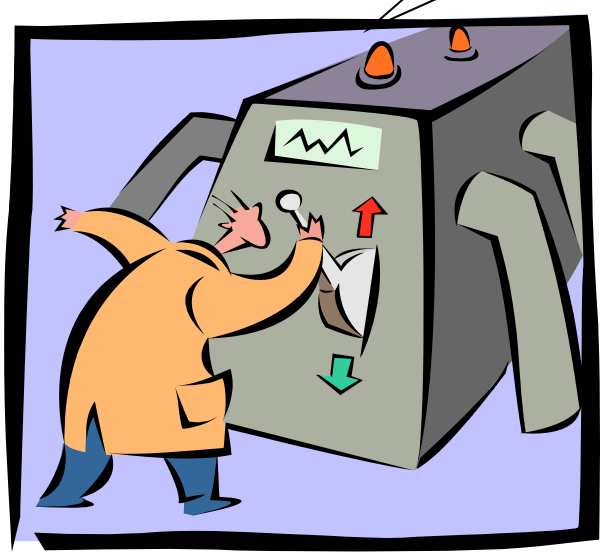
Control System

Takes human element out of the equation; does not depend on an operator or assembler.

Has a high capability of achieving zero defects.

Machine stops when an irregularity is detected.

"There must have been an error detected; the machine shut down by itself!"



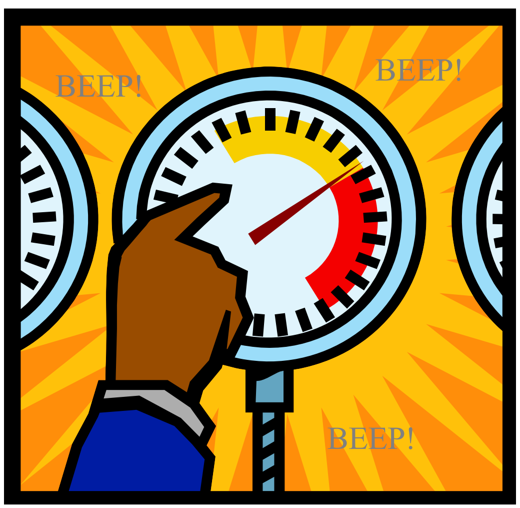
Warning System

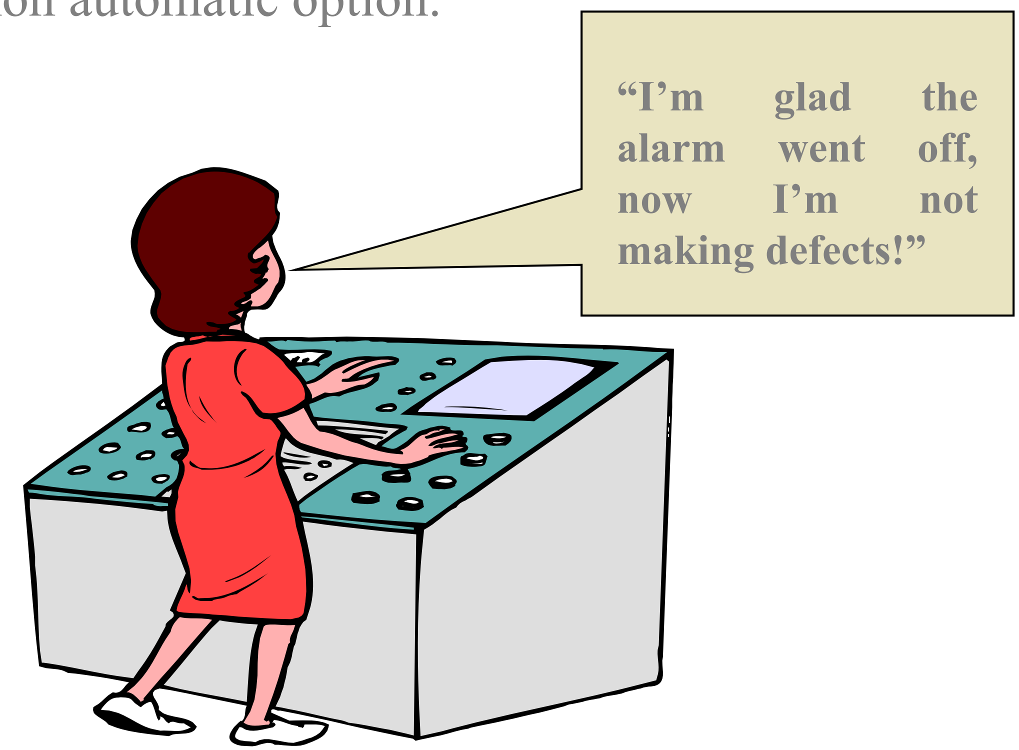
Sometimes an automatic shut off system is not an option.

A warning or alarm system can be used to get an operators attention.

Below left is an example of an alarm system using dials, lights and sounds to bring attention to the problem.

Color coding is also an effective non automatic option.

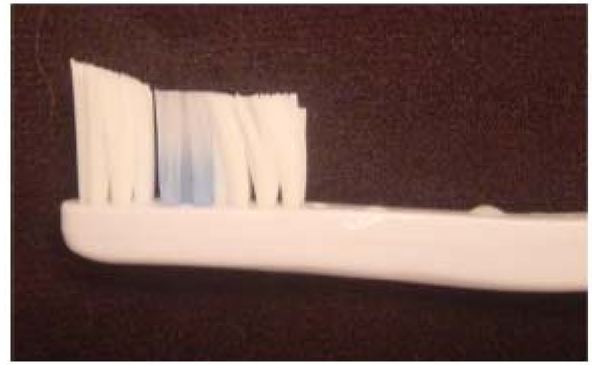




Poka Yoke Example

Mistake Proofing for cows

Poka Yoke Example



When the forces of commercial self-interest and of dental hygiene combine, how can it not lead to mistake-proofing?
This toothbrush has colored bristles that become clear at the tips of the bristles through use. When it starts to look like the brush on the right, it is time to buy a new toothbrush.
Planned obsolescence at its best.

Real World Examples of Poka-Yoke Devices

PREVENTION

- Fuelling area of car has three mistake-proofing devices:
  - filling pipe insert keeps larger leaded-fuel nozzle from being inserted
  - gas cap tether does not allow the motorist to drive off without the cap
  - gas cap is fitted with ratchet to signal proper tightness and prevent over-tightening.

Statistical Process Control (SPC)

- SPC was developed by Walter A. Shewhart in 1924
- Historically, SPC has been used to monitor & control output ‘Y’
- In DMAIC, we apply SPC to control X’s (remember ‘Y’ is only monitored)
- However, sometimes applying SPC to ‘Y’ can also be beneficial in detecting trends
- About SPC
  - Aids visual monitoring & controlling
  - Depends heavily on data collection

Foundation of SPC

- It forms data into patterns which can be statistically tested and, as a result, leads to information about the behavior of process output / control variable characteristics

- It graphically represents output / control variable performance

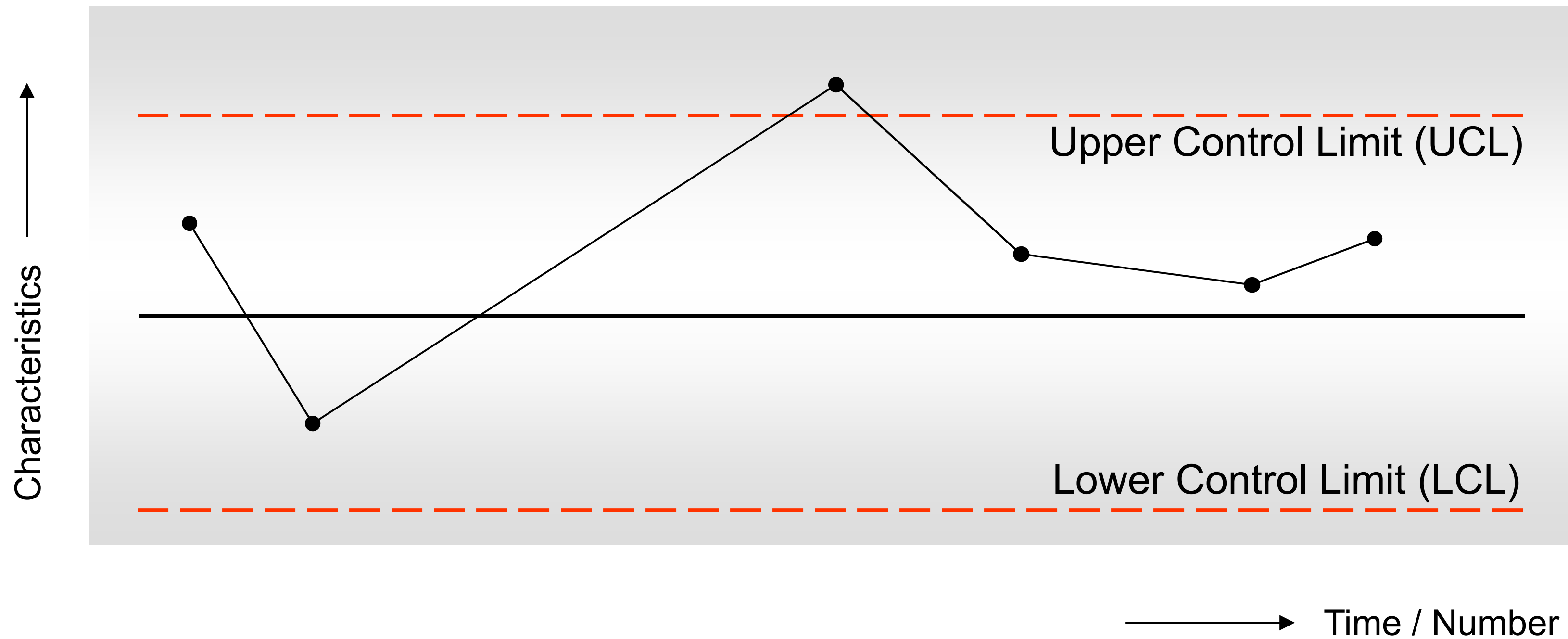
- It detects assignable causes which affects the central tendency and/or variability of the cause system

- It serves as a probability-based decision making tool

- It points out where action can be taken with known degrees of risk and confidence

SPC Tools

- **SPC primarily uses 'Control Charts'**



- 'Process Control' is inherent to process characteristics as against 'Process Capability' which is measured as per outside targets & specifications

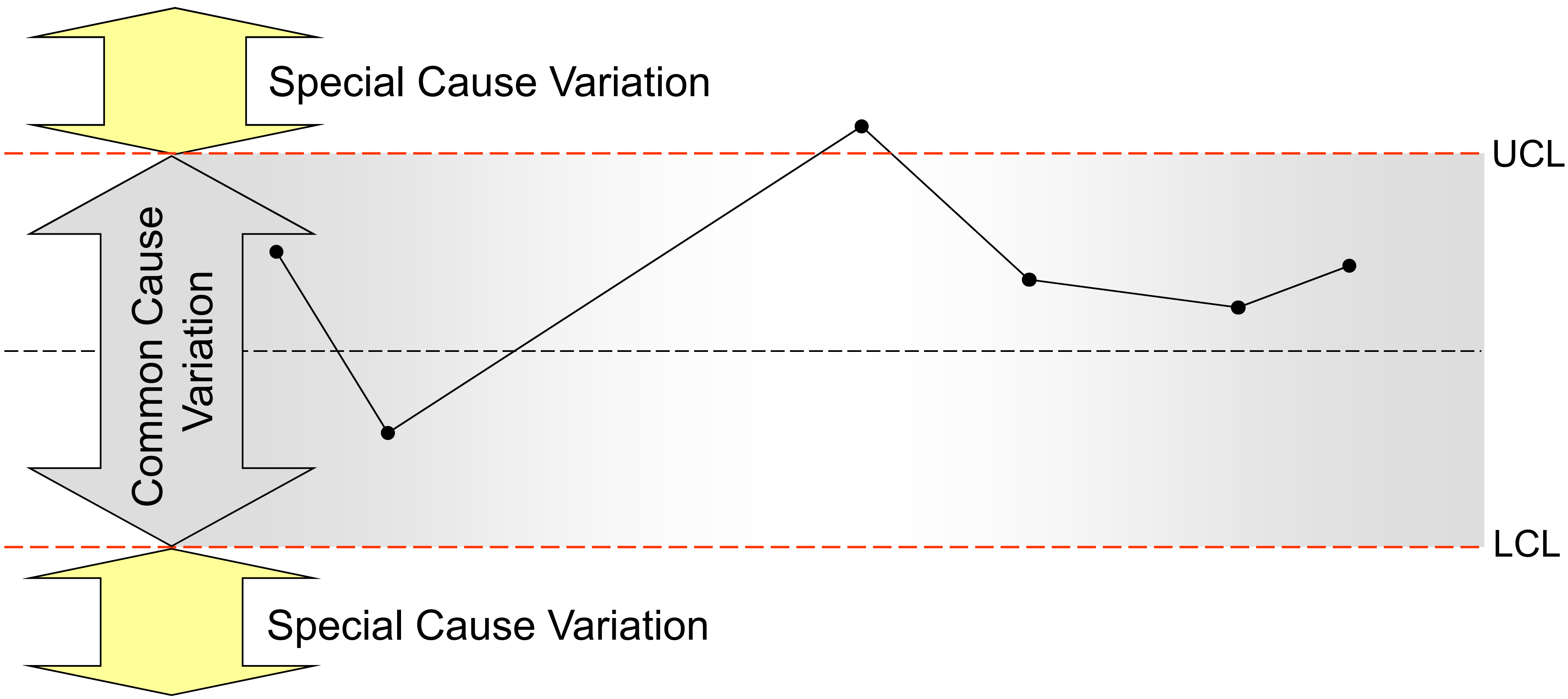
Basics of Control Charts

- Control charts are useful for tracking process statistics over time and detecting the presence of special causes

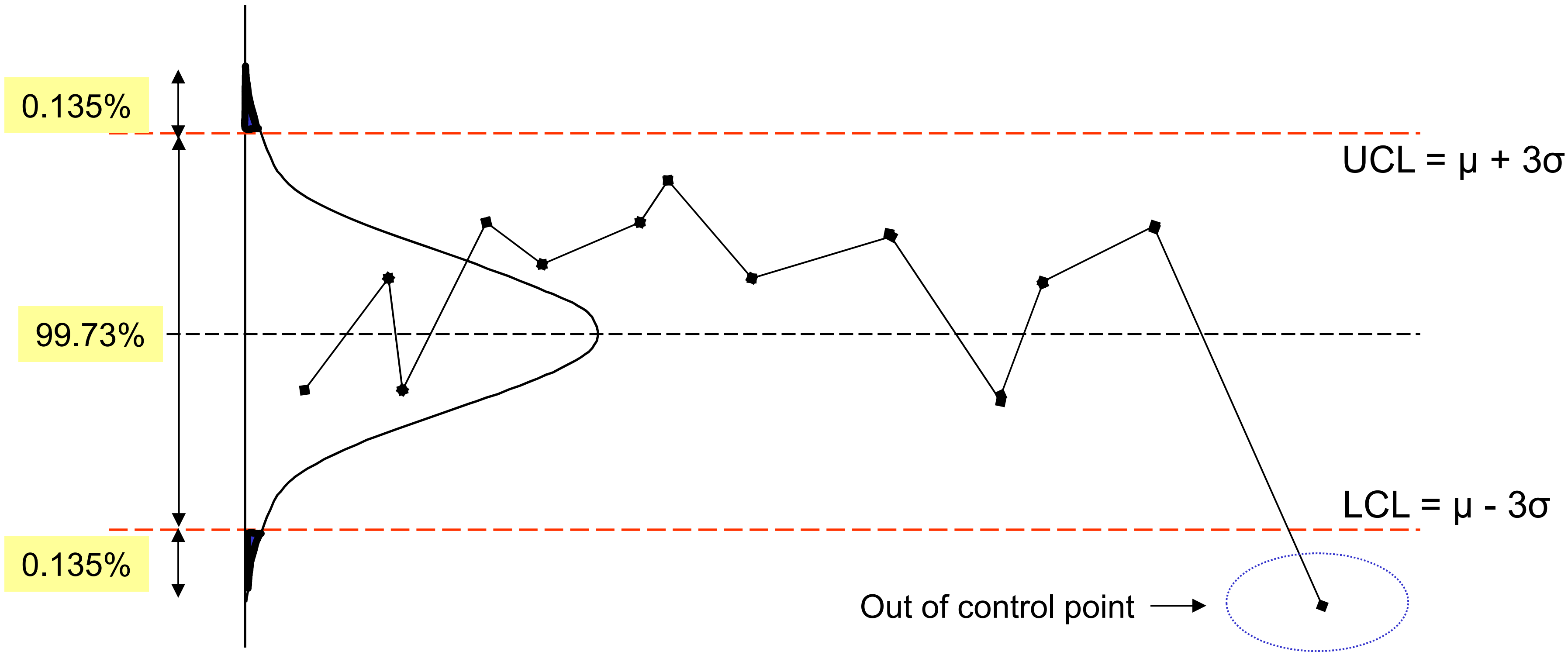
- A process statistic, such as a subgroup mean, individual observation, or weighted statistic, is plotted versus sample number or time. A "center line" is drawn at the average of the statistic being plotted for the time being charted. Two other lines—the upper and lower control limits—are drawn, by default, 3σ above and below the center line

- A process is in control when most of the points fall within the bounds of the control limits, and the points do not display any nonrandom patterns

Purpose of Control Limits



Purpose of Control Limits



Control Limits define a probabilistic level of occurrence of an 'out of control' point

Setting the Control Limits

- A standard control chart uses control limits at three standard deviations form the data mean. The probability of an out-of-control point when the process has *not* changed is only 0.27%

- If control limits are set at two standard deviations, it increases the chance of type I error

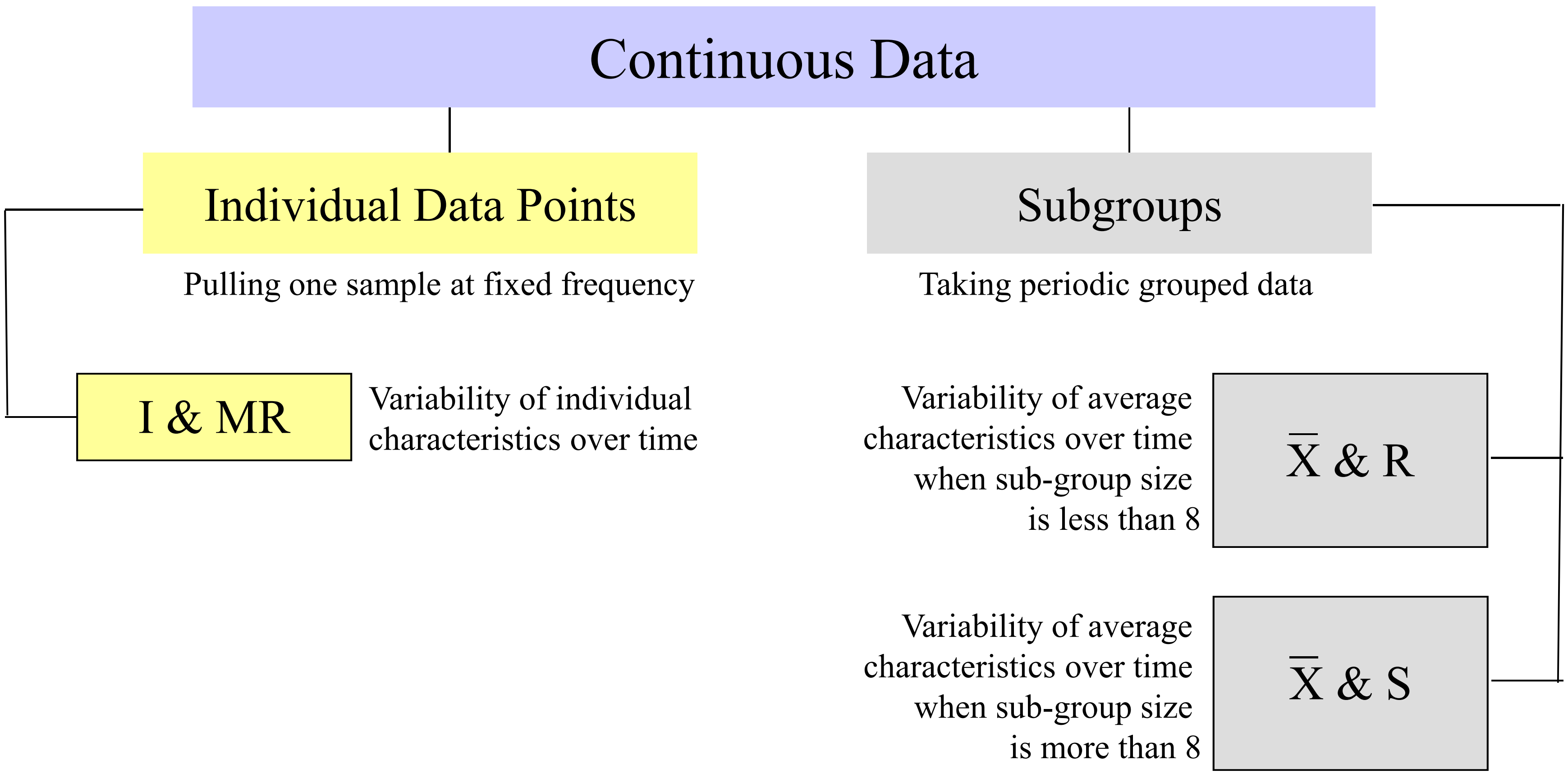
- If control limits are set at four standard deviations, it increases the chance of a type II error

- Control chart should keep in mind both type I & type II errors

Top Nine Indications of an Out of Control Process

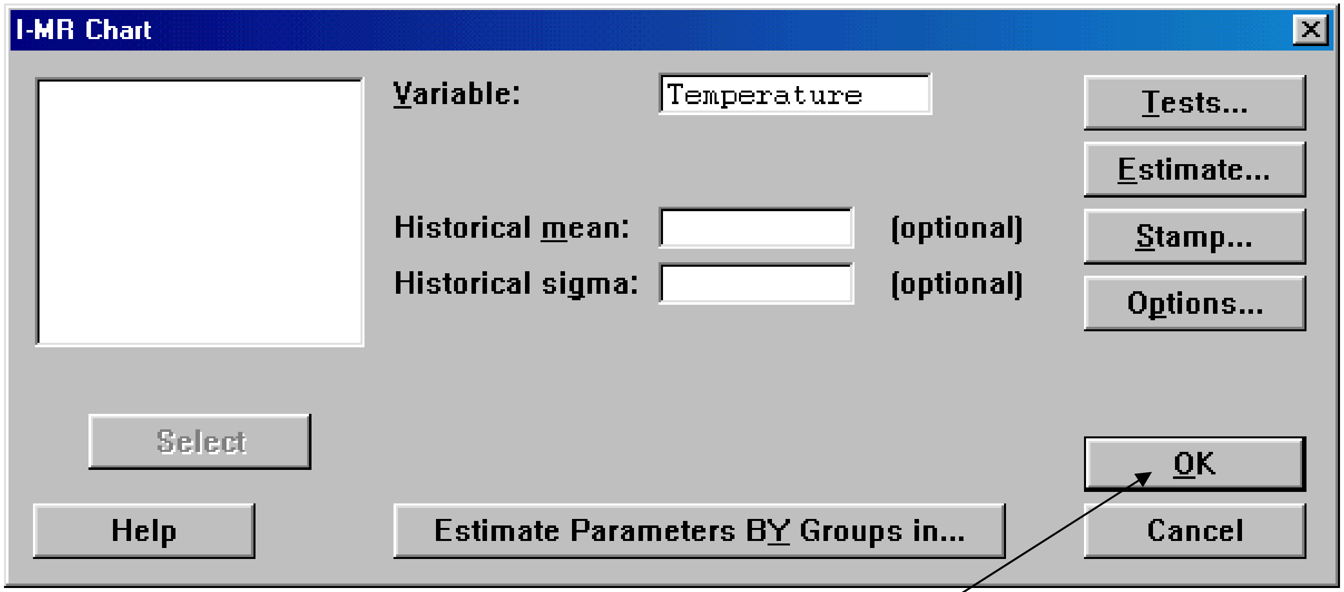
- A single point outside control limits
- Two out of three successive points between 2σ and 3σ on the same side of the centerline
- Seven successive points on the same side of the centerline
- Nine out of ten successive points on the same side of the centerline
- Twelve out of fourteen successive points on the same side of the centerline
- Consistent increase or decrease in levels
- Fourteen points alternating up and down
- Four out of five successful points beyond 1σ on the same side of the centerline
- Eight points in a row with none between $\pm 1\sigma$

Choosing An Appropriate Control Chart



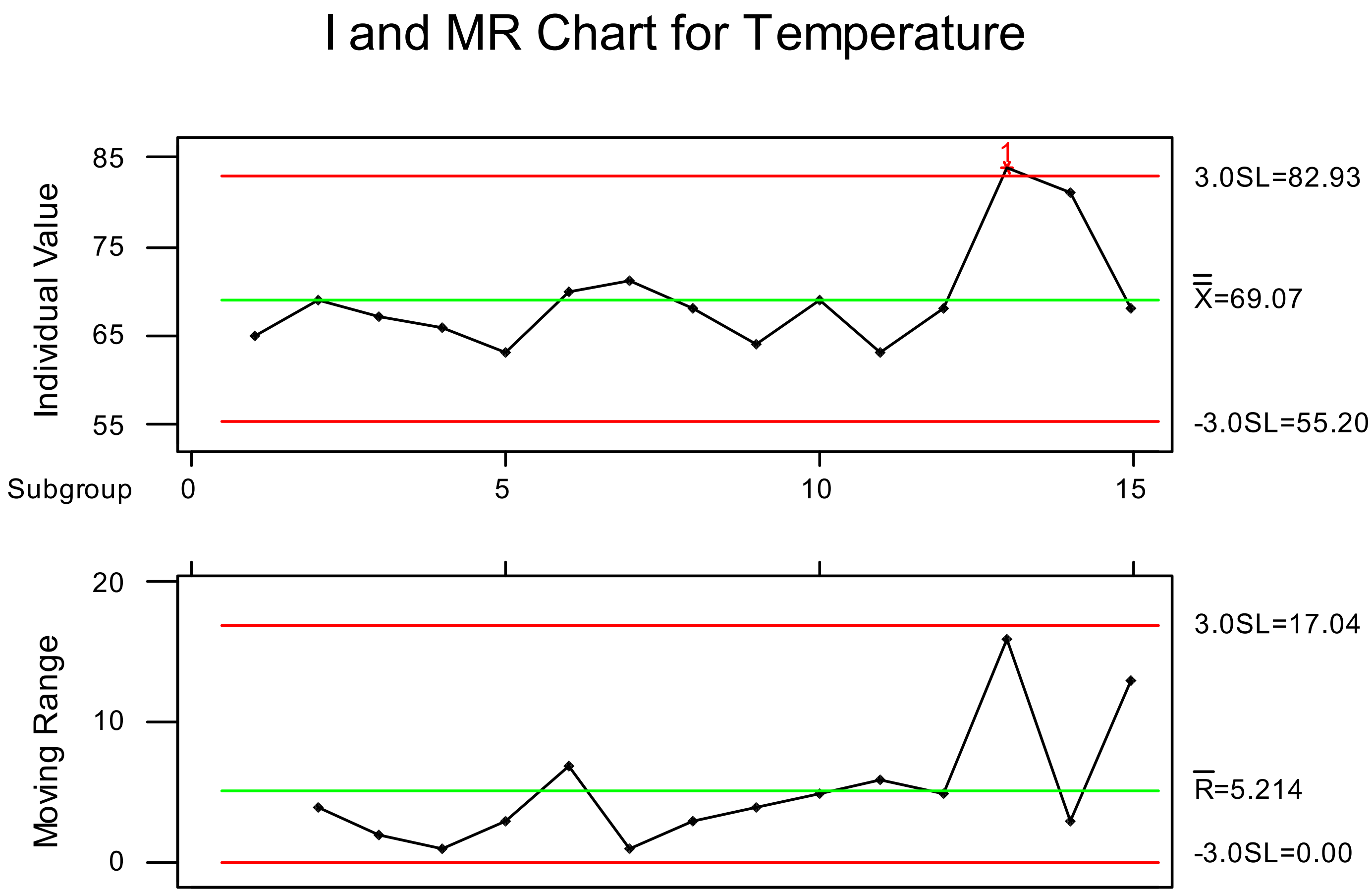
I & MR Control Chart

- STAT > CONTROL CHARTS > I-MR



Click on OK

I & MR Control Chart

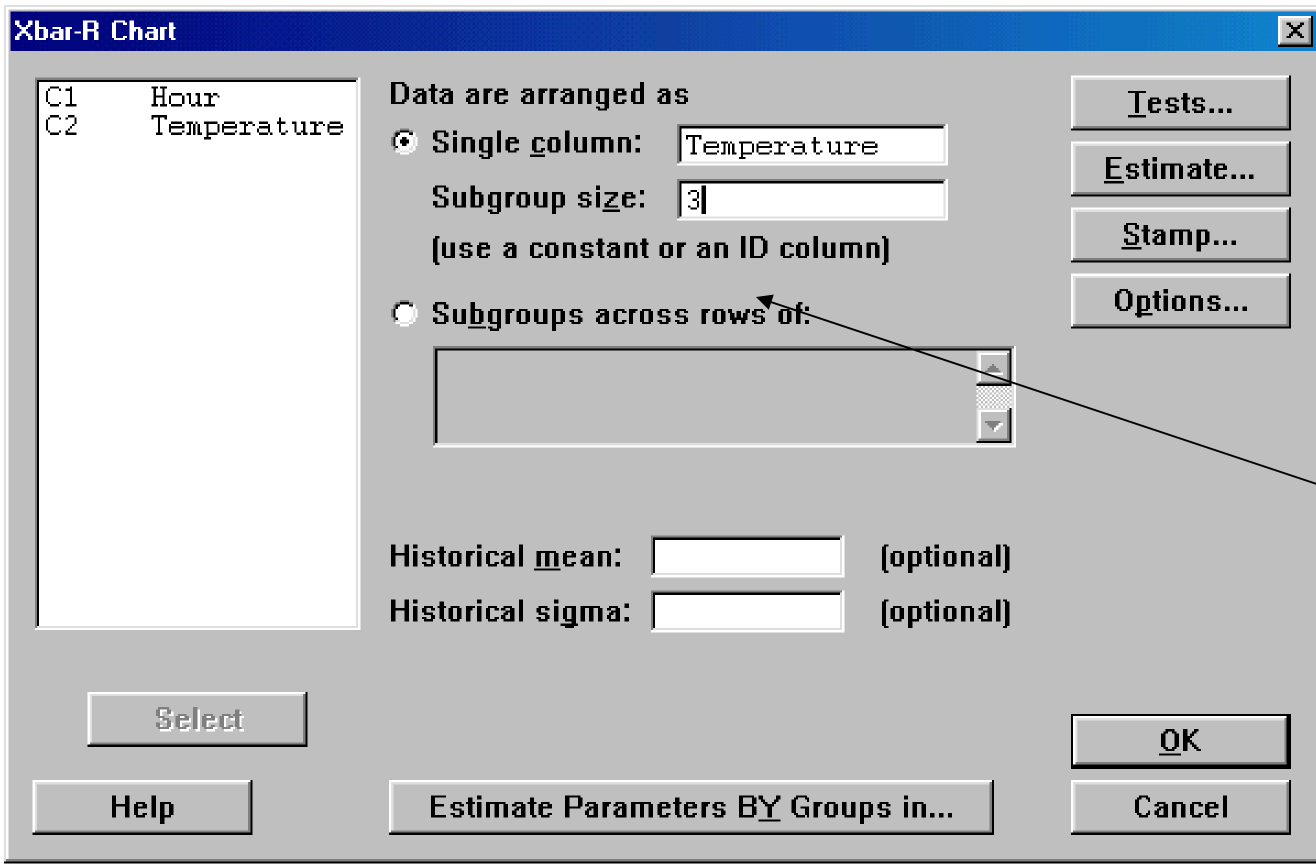


X & R Control Chart

- Let’s take data of the previous example only. Assume that the data on temperature was collected using three different probes & below table gives three readings per hour, each for one probe, over 5 hours (5 samples, each of sub-group size 3)

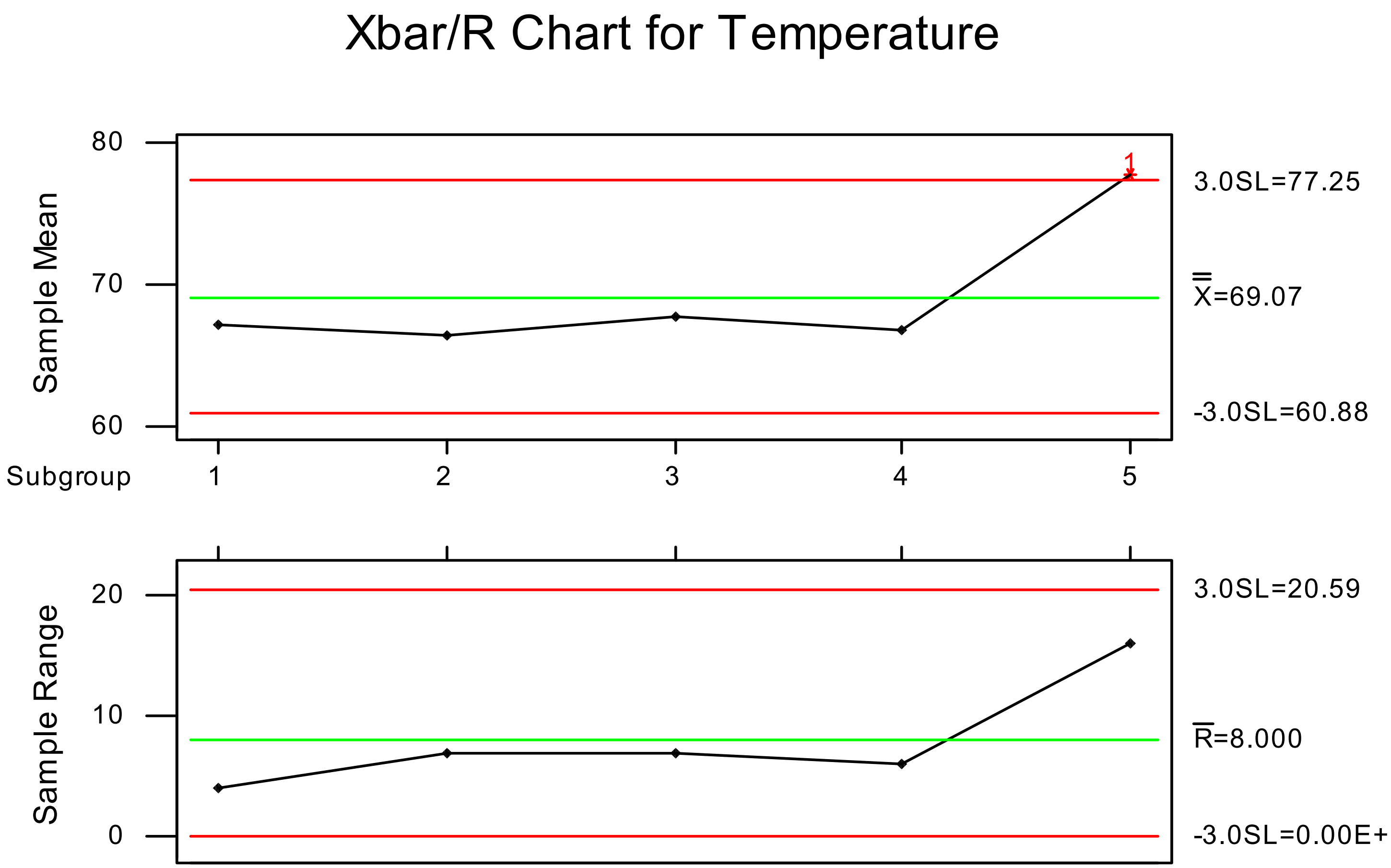
Hour	Temperature
1	65
1	69
1	67
2	66
2	63
2	70
3	71
3	68
3	64
4	69
4	63
4	68
5	84
5	81
5	68

X & R Control Chart



A new input is now required – sub-group size

X & R Control Chart



Example

You work at an automobile engine assembly plant. One of the parts, a camshaft, must be 600 mm $\pm$2 mm long to meet engineering specifications. There has been a chronic problem with camshaft length being out of specification, which causes poor-fitting assemblies, resulting in high scrap and rework rates. Your supervisor wants to run X and R charts to monitor this characteristic, so for a month, you collect a total of 100 observations (20 samples of 5 camshafts each) from all the camshafts used at the plant, and 100 observations from each of your suppliers. First you will look at camshafts produced by Supplier 2.

X & S Control Chart

- Data is collected on number of runs scored by a cricketer sub-grouped by rival countries which is a vital 'X' in winning. Sub-group size is sufficiently large & is varying since number of matches played against each country is not same.

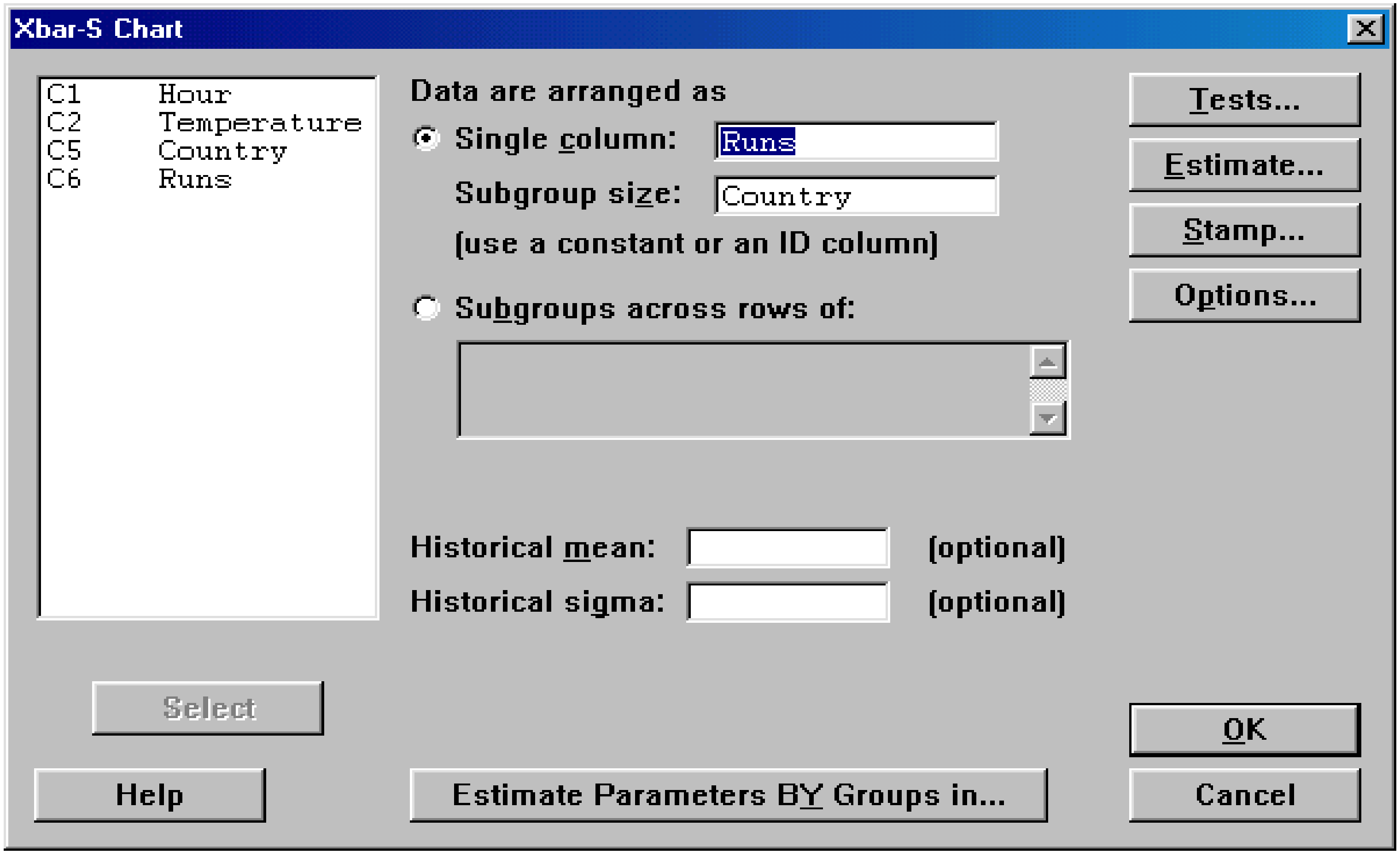
Country	Runs in a Match
1	65
1	6
1	0
1	19
1	63
1	112
1	12
1	35
2	5
2	9
2	63
2	32
2	98
2	81
2	28
2	9
2	16
2	41

Country	Runs in a Match
3	54
3	32
3	69
3	89
3	12
3	3
3	116
3	21
3	26
3	65
4	0
4	3
4	6
4	15
4	63
4	32
4	24
4	16

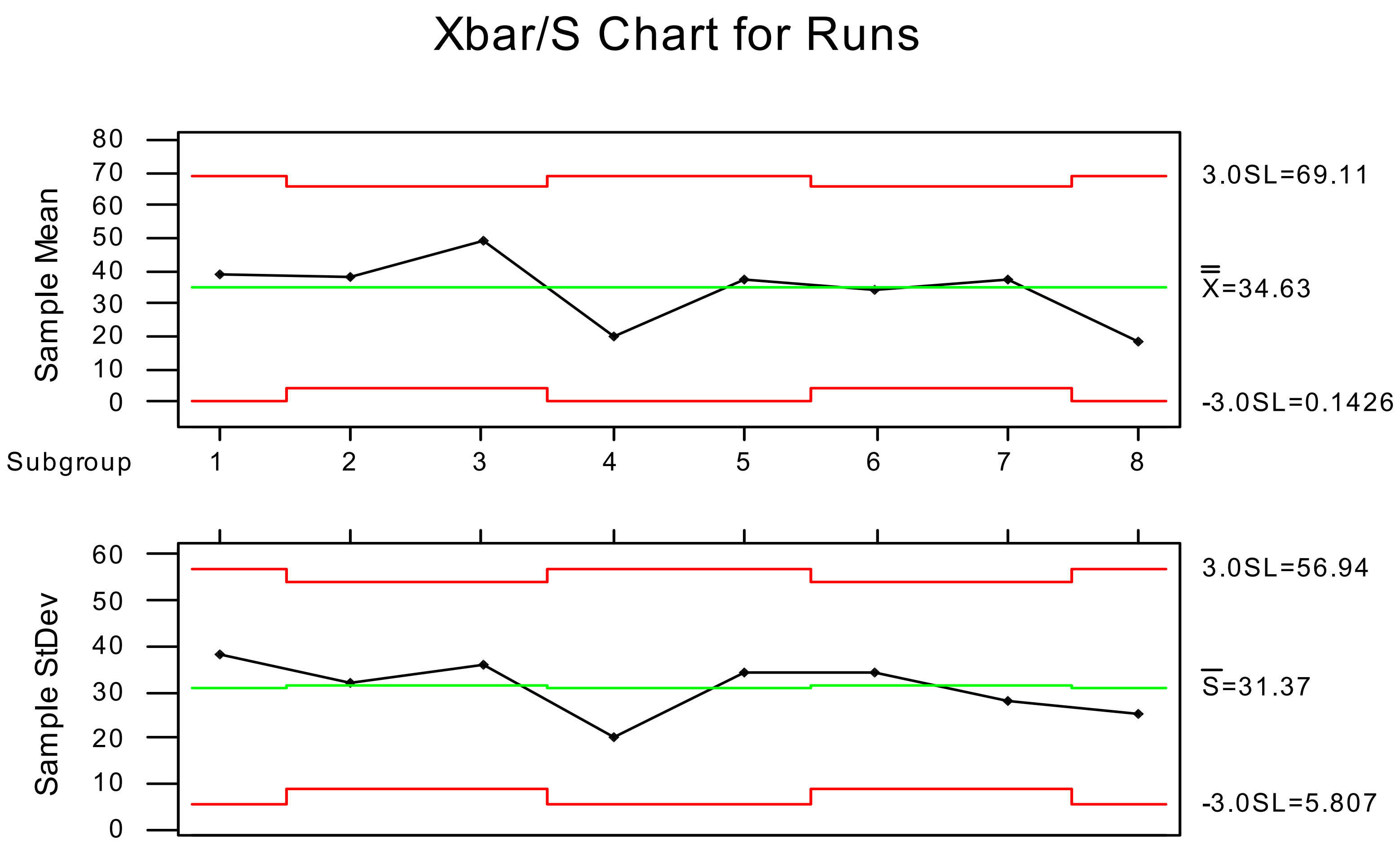
Country	Runs in a Match
5	62
5	3
5	9
5	21
5	60
5	101
5	9
5	32
6	2
6	6
6	60
6	29
6	95
6	81
6	25
6	3
6	0
6	36

Country	Runs in a Match
7	32
7	29
7	9
7	33
7	9
7	101
7	56
7	18
7	23
7	62
8	6
8	0
8	3
8	65
8	52
8	3
8	0
8	18

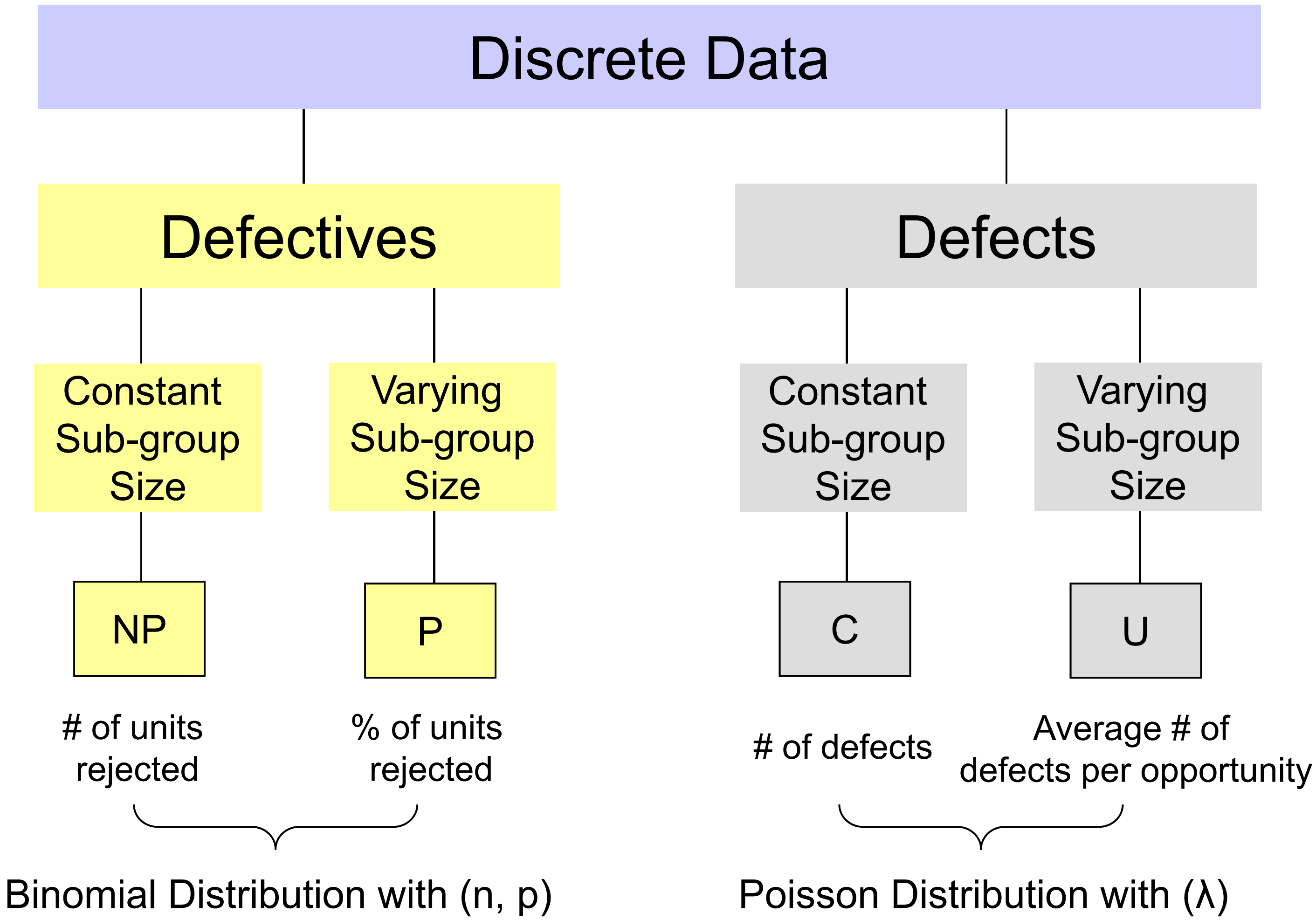
X & S Control Chart



X & S Control Chart



Choosing An Appropriate Control Chart

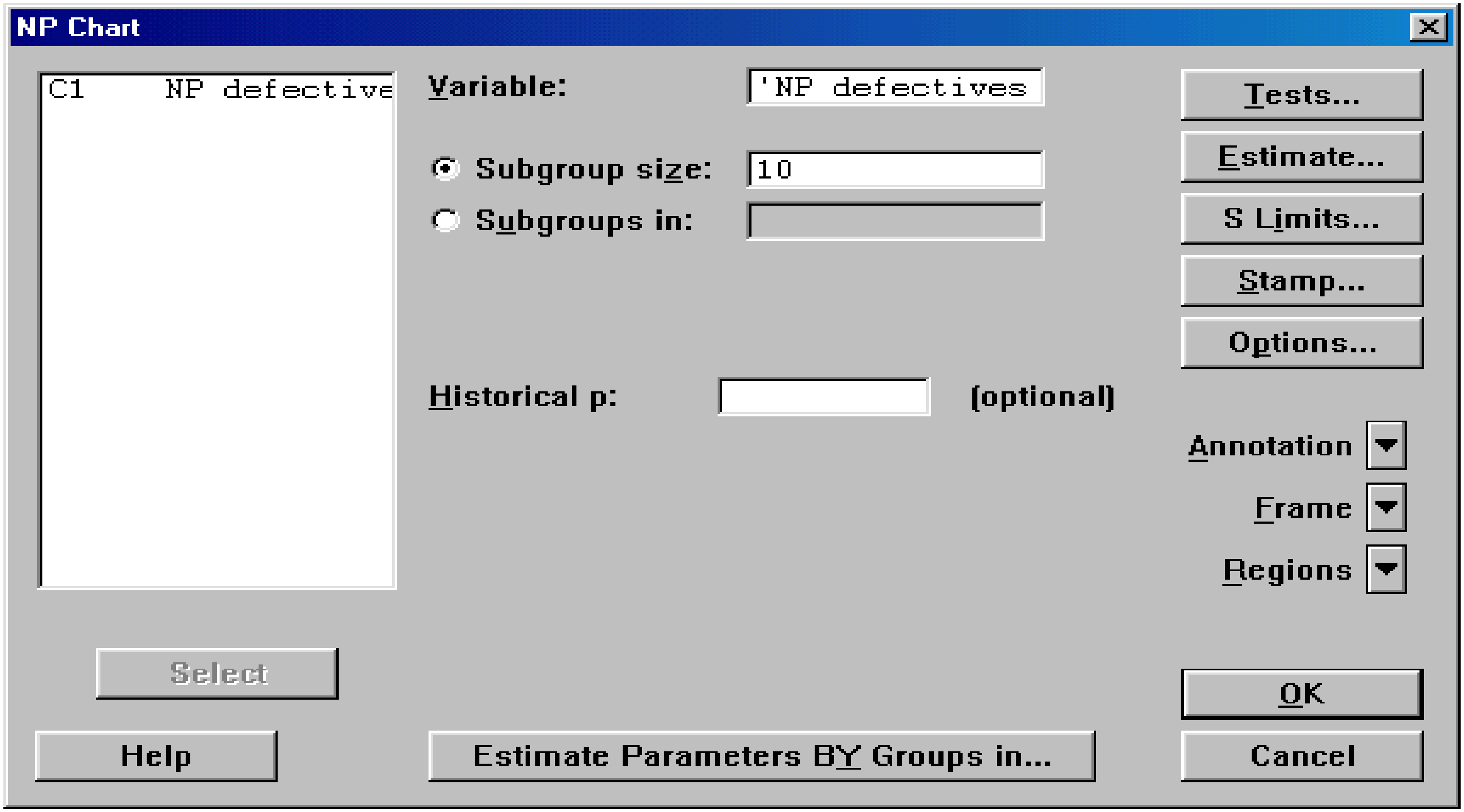


NP Example

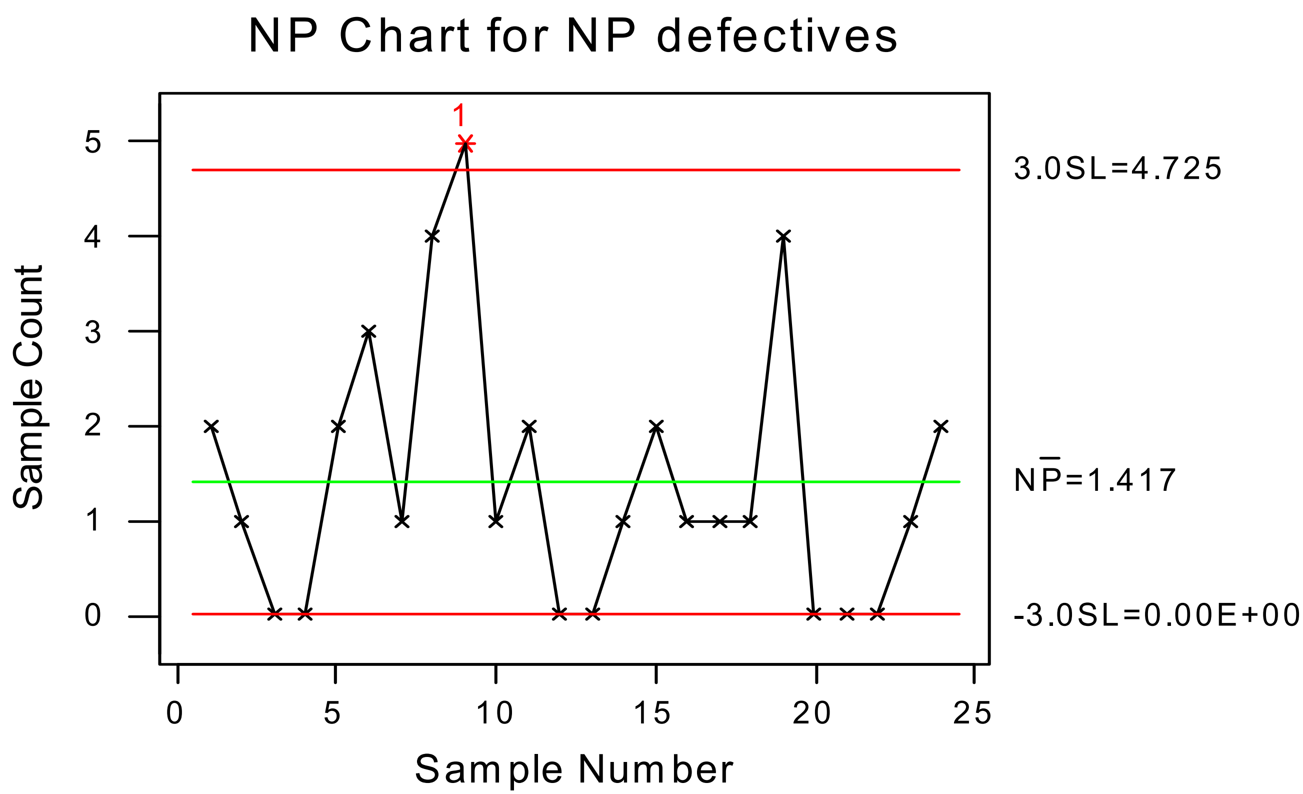
- Let's assume that the quality control department checks the quality of finished goods by sampling a batch of 10 items from the produced lot every hour. If items are found out of control limits consistently in any given day, production process has to be stopped for the next day. They collect the following data over 24 hours:

Hour	Defectives	Hour	Defectives
1	2	13	0
2	1	14	1
3	0	15	2
4	0	16	1
5	2	17	1
6	3	18	1
7	1	19	4
8	4	20	0
9	5	21	0
10	1	22	0
11	2	23	1
12	0	24	2

NP Control Chart



NP Control Chart



Example

You work in a toy manufacturing company and your job is to inspect the number of defective toy. You inspect 200 samples in each lot and then decide to create an NP chart to monitor the number of defectives.

P Control Chart

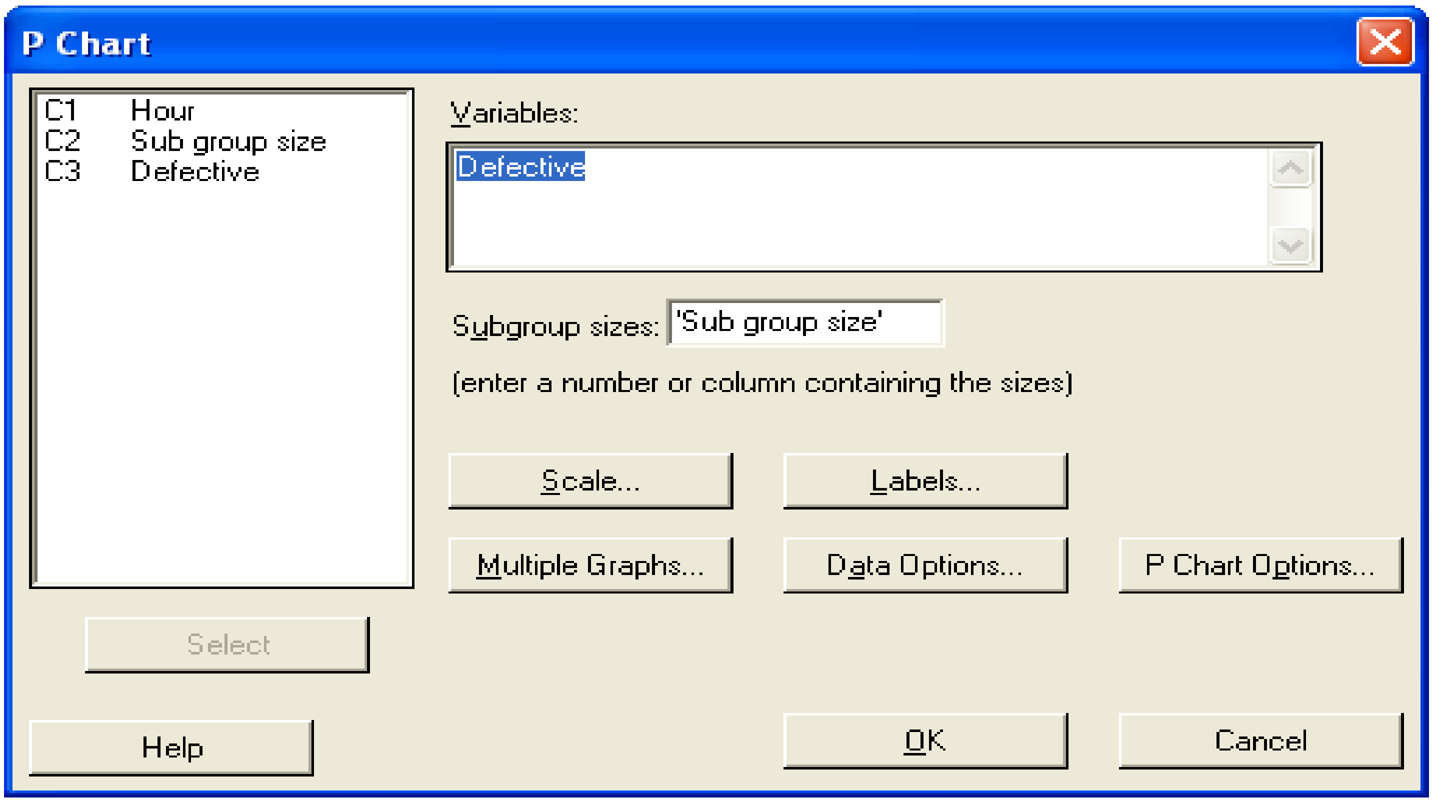
- Now let's vary the sub-group size, i.e. number of items tested for defectiveness varies from hour-to-hour

Hour	Sub-group-size	Defectives
1	10	2
2	10	1
3	10	0
4	20	0
5	10	2
6	20	3
7	10	1
8	20	4
9	20	5
10	10	1
11	10	2
12	10	0

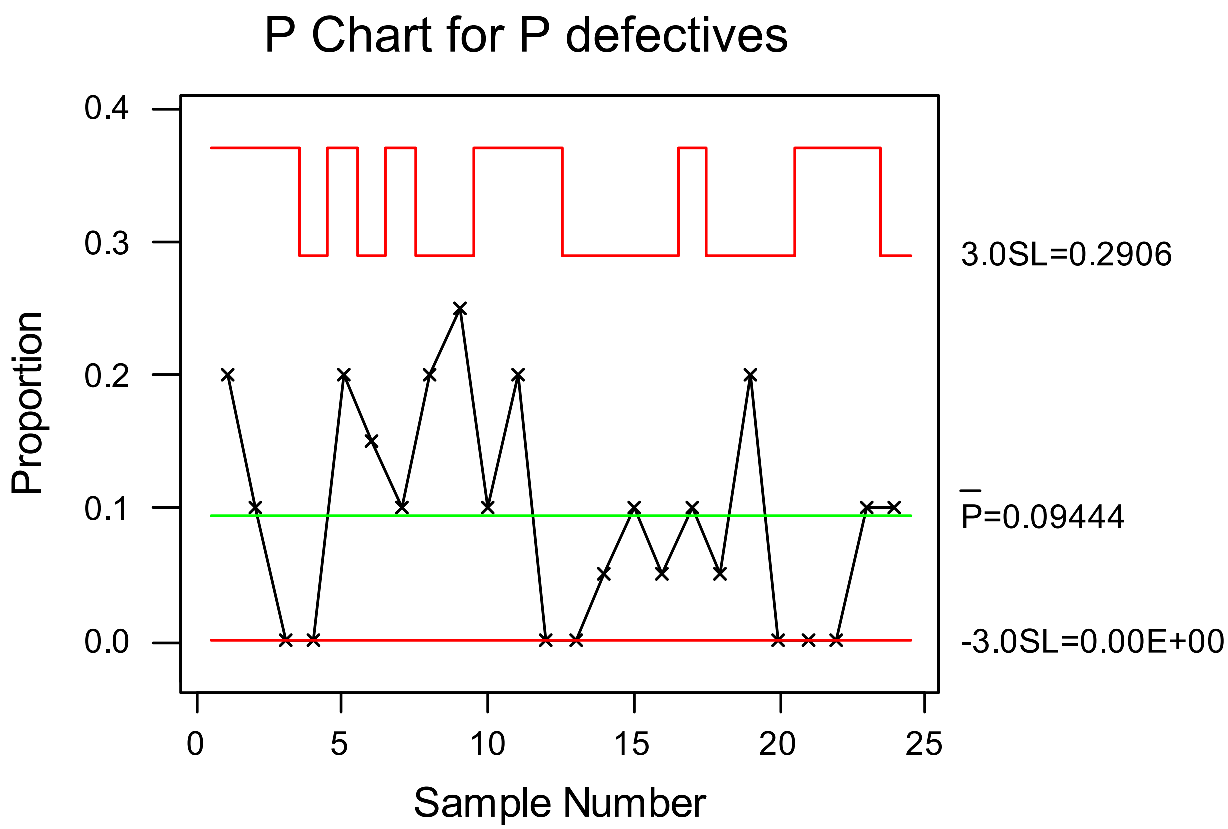
Hour	Sub-group-size	Defectives
13	20	0
14	20	1
15	20	2
16	20	1
17	10	1
18	20	1
19	20	4
20	20	0
21	10	0
22	10	0
23	10	1
24	20	2

P Control Chart

- STAT > CONTROL CHART > P



P Control Chart



Example

Suppose you work in a plant that manufactures picture tubes for televisions. For each lot, you pull some of the tubes and do a visual inspection. If a tube has scratches on the inside, you reject it.

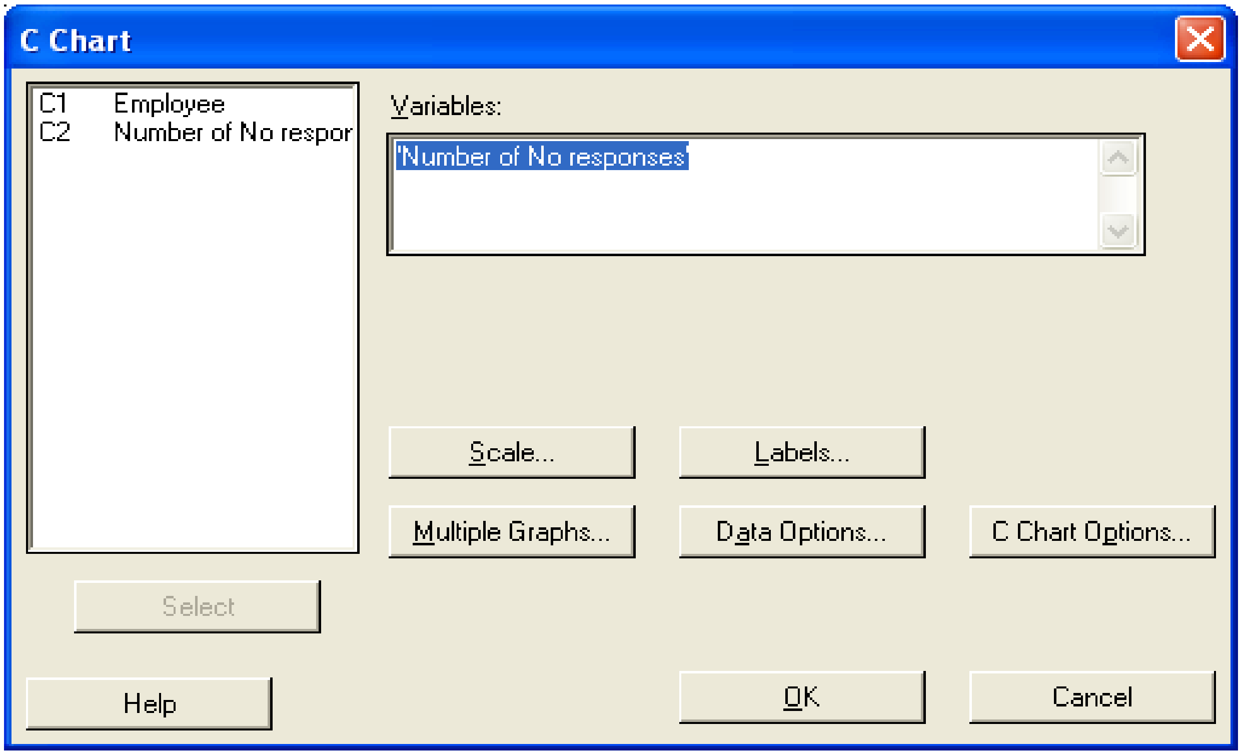
C Control Chart

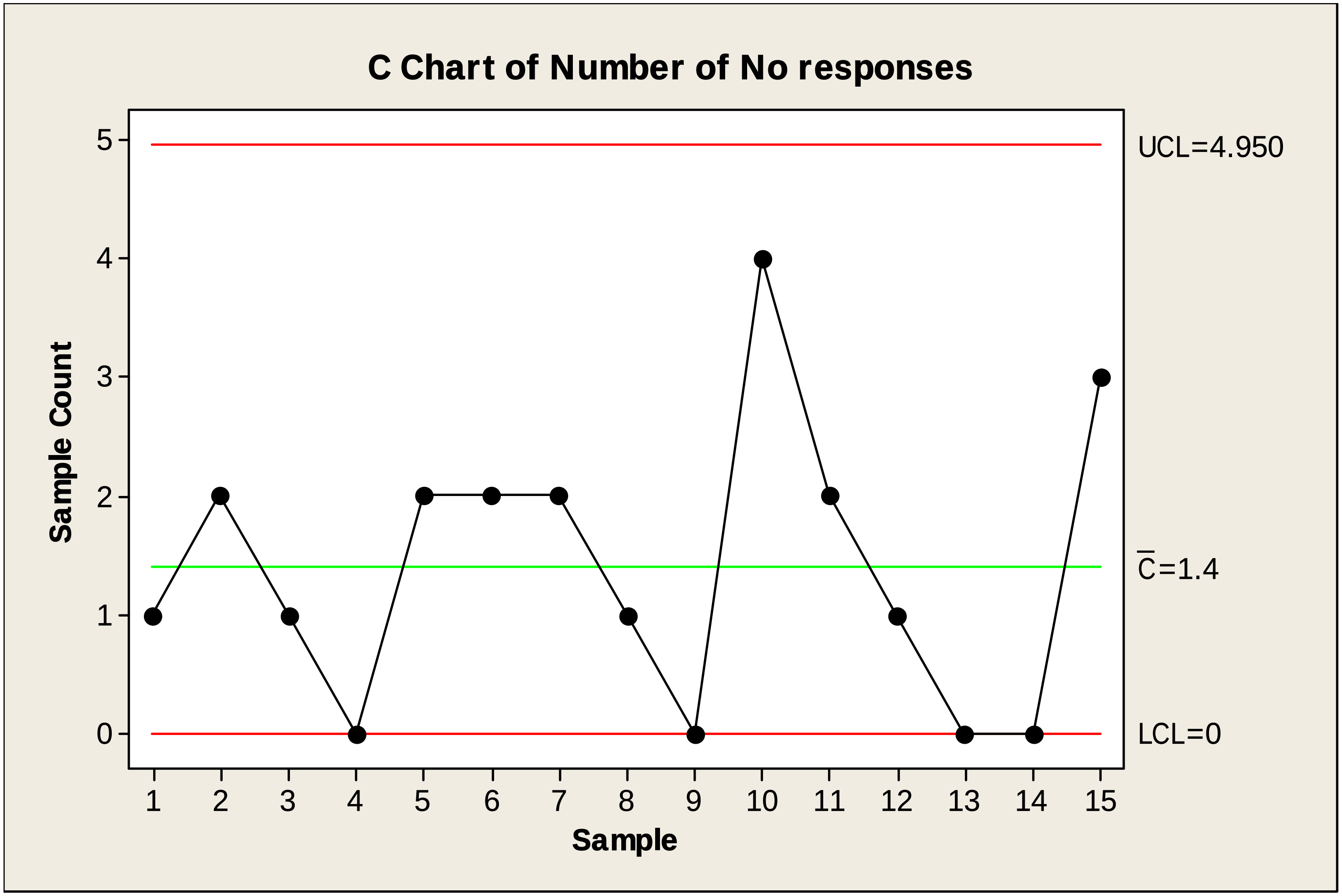
- Let's assume that the customer service department administers a questionnaire on employees which has to be answered in 'yes / no'. There are total 15 questions. Each question that is answered in a 'no' is a defect. These questions form a vital 'X' in measuring employee satisfaction.

Employee	Number of 'No' responses
1	1
2	2
3	1
4	0
5	2
6	2
7	2
8	1
9	0
10	4
11	2
12	1
13	0
14	0
15	3

C Control Chart

- STAT > CONTROL CHART > C




C Chart of Number of No responses
Sample Count
5
4
3
2
1
0
UCL=4.950
C̄=1.4
LCL=0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Sample

Example

Suppose you work for a linen manufacturer. Each 100 square yards of fabric can contain a certain number of blemishes before it is rejected. For quality purposes, you want to track the number of blemishes per 100 square yards over a period of several days, to see if your process is behaving predictably.

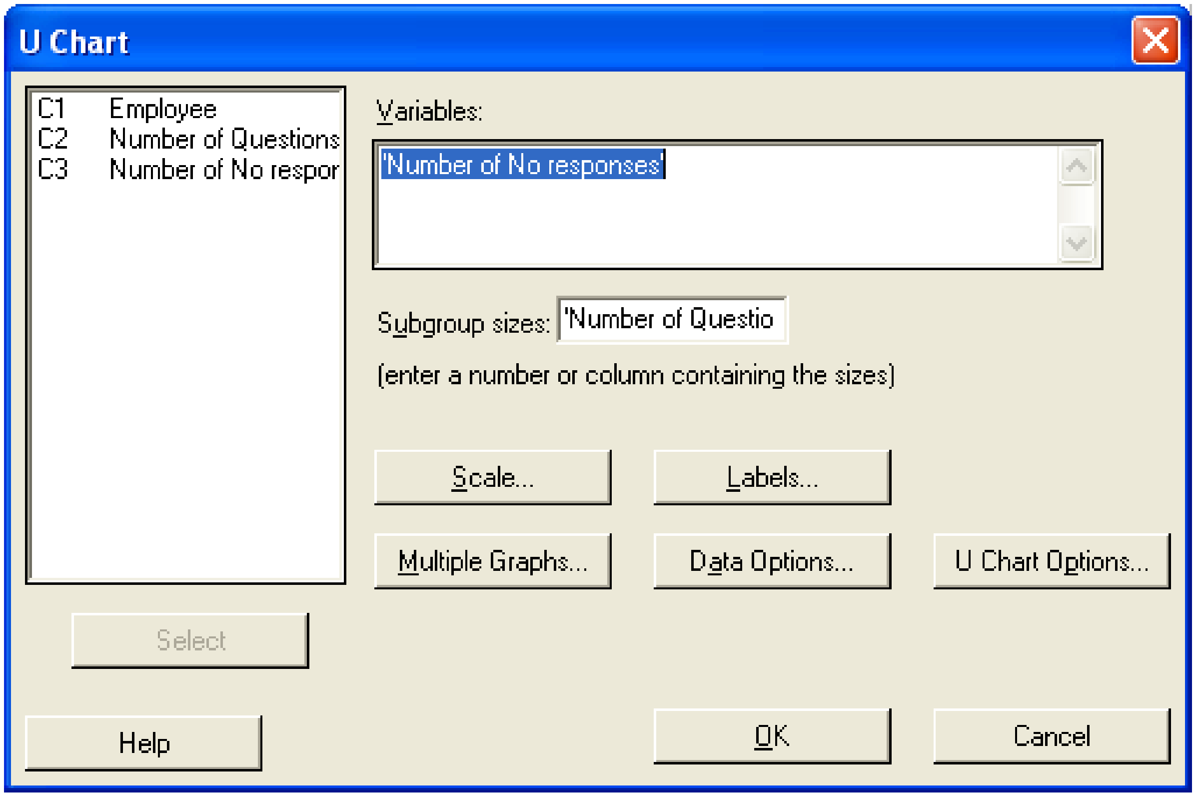
U Control Chart

- Let's slightly change the data used in example of C chart. Let's assume that the customer service department now administers two questionnaires on employees, one with 10 & another with 20 questions, i.e. sub-group size varies. They have to be answered in 'yes / no'. Each question that is answered in a 'no' is a defect.

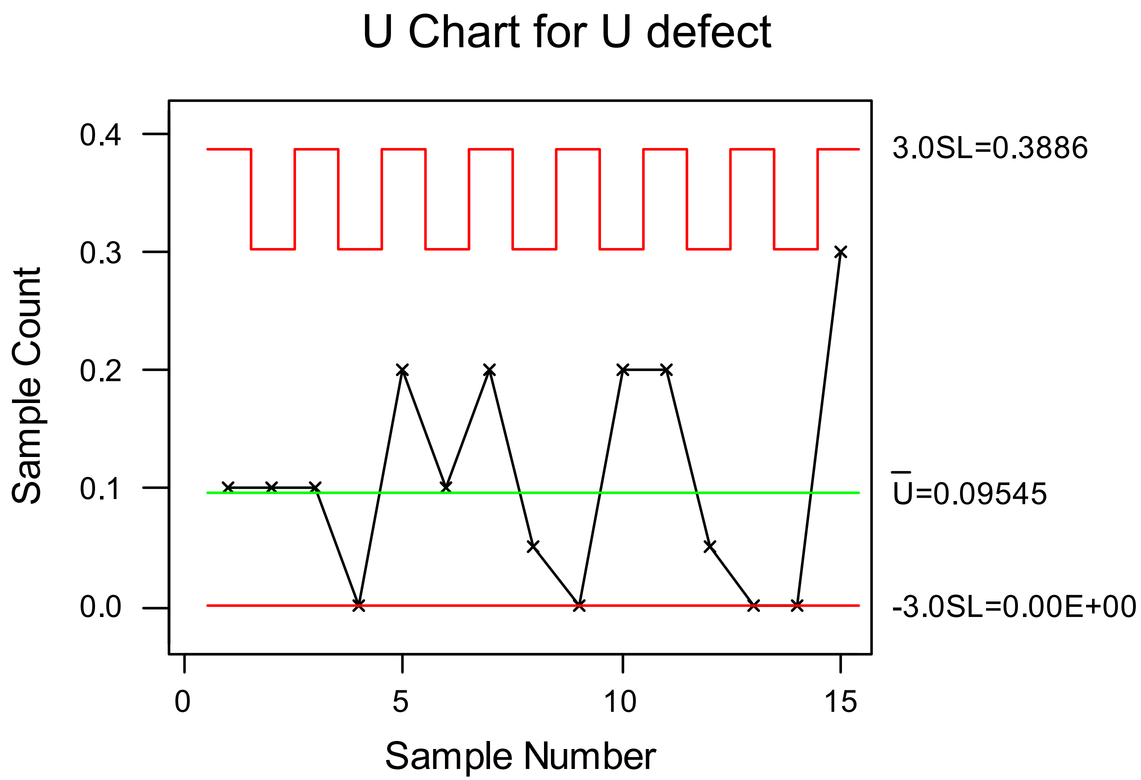
Employee	Number of Questions	Number of 'No' responses
1	10	1
2	20	2
3	10	1
4	20	0
5	10	2
6	20	2
7	10	2
8	20	1
9	10	0
10	20	4
11	10	2
12	20	1
13	10	0
14	20	0
15	10	3

U Control Chart

- STAT > CONTROL CHART > U



U Control Chart



Specification Limits v/s Control Limits

Specification Limits

- Come from Engineering or customer requirements
- Represents what someone wants a process to do
- Can sometimes be changed by changing the requirements of the product or service

Control Limits

- Come from calculations on the process data
- Represents what a process is actually capable of doing
- Can be changed by changing the process